

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2003, 18:54:25 ; Search time 48 Seconds
(without alignments)
3211.895 Million cell updates/sec

Title: US-09-974-973-2

Perfect score: 5865

Sequence: 1 MTAITLGLLLGIITLVST.....RVVPAATKVEGDLIVVVS 1157

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5865	100.0	1157	AAU98050	Corynebacterium fe
2	5779	98.5	1140	AAU98052	Corynebacterium mu
3	5764	98.3	1141	AAU98052	Pyruvate carboxyla
4	5759	98.2	1140	AAU98051	C glutamicum prote
5	5759	98.2	1140	AAU98051	Corynebacterium gl
6	5759	98.2	1140	AAU98053	Corynebacterium wi
7	5753	98.1	1140	AAU93971	C. glutamicum pyru
8	5751	98.1	1140	AAU93971	C glutamicum prote
9	5299.5	90.4	1139	AAU93180	Corynebacterium th
10	2678	45.7	532	AAU979302	Corynebacterium gl

11	2530	43.1	1148	22	AAU00511	Bacillus subtilis
12	2494	42.5	1146	23	ABU47612	Listeria monocytog
13	2489	42.4	1147	22	AAU33972	Staphylococcus aur
14	2465	42.0	1151	22	AGS2677	S. epidermidis ope
15	2465	42.0	1154	23	ABP38583	Staphylococcus epi
16	2460.5	42.0	1142	22	AAU35213	Enterococcus faeca
17	2441	41.6	1181	22	ABU58211	Drosophila melanog
18	2441	41.6	1181	22	ABU66604	Drosophila melanog
19	2441	41.6	1181	22	ABU66605	Drosophila melanog
20	2433.5	41.5	1196	22	ABU67309	Lactococcus lactis
21	2412	41.1	1136	23	ABU5980	Corynebacterium gl
22	2336	39.8	461	22	AAU79303	Staphylococcus aur
23	2300	39.2	1073	22	AAU36768	Corynebacterium gl
24	1546.5	26.4	320	22	AAU79300	Corynebacterium gl
25	1546.5	26.4	320	22	AAU79301	Corynebacterium gl
26	1371	23.4	272	22	AAU79298	Corynebacterium gl
27	1371	23.4	272	22	AAU79299	Corynebacterium gl
28	1032.5	17.6	447	17	AAU05207	Anabaena biotin bi
29	1032.5	17.6	447	19	AAU70402	Anabaena biotin ca
30	1032.5	17.6	448	15	AAU51080	Biotin carboxylase
31	1005.5	17.1	453	15	AAU51083	fabG gene encoding
32	1005.5	17.1	453	17	AAU05208	Synechococcus biot
33	1005.5	17.1	453	19	AAU70403	Synechococcus biot
34	980	16.7	425	21	AAU11663	A. vitis hyperpens
35	979.5	16.7	471	22	AAU33719	Pseudomonas aerugi
36	956	16.3	449	22	AAU38292	Salmonella typhi c
37	954.5	16.3	455	23	ABU54091	Lactococcus lactis
38	950	16.2	448	22	AAU35567	Haemophilus influe
39	947.5	16.2	456	23	ABU28017	Streptococcus poly
40	940	16.0	449	16	AAU66742	Biotin-carboxylase
41	940	16.0	449	22	AAU34737	E. coli cellular p
42	937.5	16.0	454	20	AAU34779	Chlamydia pneumoni
43	936.5	16.0	455	22	AAU38024	Streptococcus pneu
44	935.5	16.0	455	22	AAU37771	Streptococcus pneu
45	933.5	15.9	455	22	AAU01054	CPE 57 protein seq

ALIGNMENTS

RESULT 1
AAU98050
ID AAU98050 standard; Protein; 1157 AA.
XX
AC AAU98050;
XX
DT 27-AUG-2002 (first entry)
XX
DE Corynebacterium feedback-resistant pyruvate carboxylase enzyme.
XX
KW Feedback-resistant; pyruvate carboxylase; enzyme;
XX aspartic acid feedback inhibition resistant.
XX
OS Corynebacterium glutamicum.
XX
PH Key Location/Qualifiers
FT Region 1..18
FT /note= "Specifically claimed in claim 18"
FT Region 164..176
FT /note= "Specifically claimed in claim 18"
FT Region 193..205
FT /note= "Specifically claimed in claim 18"
FT Region 217..229
FT /note= "Specifically claimed in claim 18"
FT Region 238..250
FT /note= "Specifically claimed in claim 18"
FT Region 466..478
FT /note= "Specifically claimed in claim 18";
XX WO200231158-A2.
XX 18-APR-2002.
XX

PF 12-OCT-2001; 2001WO-US31893.
 XX 13-OCT-2000; 2000US-239913P.
 XX (ARCH) ARCHER-DANIELS MIDLAND CO.
 PA Hanke PD;
 XX WPI: 2002-463267/49.
 XX N-PSDB; ABK52832.
 DR Novel mutated, feedback resistant pyruvate carboxylase enzyme
 PT polypeptide, useful for producing amino acids e.g. L-lysine,
 PT L-threonine, L-glycine, L-glutamic acid, L-proline and L-methionine and
 PT L-isoleucine
 XX Claim 17: Fig 1; 42pp; English.
 XX The present invention relates to a new mutated, feedback-resistant
 CC pyruvate carboxylase enzyme. The invention is useful for producing an
 CC amino acid (e.g. L-lys, L-thr, L-Met, L-Ile, L-Glu, L-Arg and L-Pro),
 CC by culturing a host cell in a suitable media, and separating the amino
 CC acid from the medium. The vector of the invention is useful for
 CC replacement of a wild-type pyruvate carboxylase gene, with a feedback
 CC resistant pyruvate carboxylase gene, in a Corynebacterium glutamicum, by
 CC replacing a genomic copy of the wild-type pyruvate carboxylase gene with
 CC a selectable marker gene through homologous recombination to form a first
 CC recombinant strain, and replacing the selectable marker gene in the
 CC first recombinant strain, with feedback resistant pyruvate carboxylase
 CC gene through homologous recombination to form a second recombinant
 CC strain, where the homologous recombination in the above steps, occurs
 CC between the host cell and the vector. The feedback-resistant pyruvate
 CC carboxylase enzyme is resistant to feedback inhibition from aspartic
 CC acid. The present amino acid sequence represents the feedback-resistant
 CC pyruvate carboxylase enzyme of the invention.
 XX Sequence 1157 AA;
 SQ Query Match 100.0%; Score 5865; DB 23; Length 1157;
 Best Local Similarity 100.0%; Pred. NO. 0;
 Matches 1157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MTATLGLLLKGIITLVSTSTSLPAFKKILVANRGEIAVRAFAAETGAATVAIYP 60
 DB 1 MTATLGLLLKGIITLVSTSTSLPAFKKILVANRGEIAVRAFAAETGAATVAIYP 60
 QY 61 REDRGSFHRSFASAVRIGTSGSPVKAYLDIDEIIGAACKVKADAIYPGYGFLSENAOLA 120
 DB 61 REDRGSFHRSFASAVRIGTSGSPVKAYLDIDEIIGAACKVKADAIYPGYGFLSENAOLA 120
 QY 121 RECAENGITFTGPTPEVLDLTDGKSRVATAAKKAGLPVLAESTPSKNIDDIVKSAEGQTY 180
 DB 121 RECAENGITFTGPTPEVLDLTDGKSRVATAAKKAGLPVLAESTPSKNIDDIVKSAEGQTY 180
 QY 181 PIFVKAVAGGGGRMRFSVPDEURKLATEASREAEAFGDSGYVERAVINPOHIEVQI 240
 DB 181 PIFVKAVAGGGGRMRFSVPDEURKLATEASREAEAFGDSGYVERAVINPOHIEVQI 240
 QY 241 LGDRTGEVHHLYERDCSLQRRHQKVVETAPAOHLDPRLDRICADAVKFCRSIGYQAGT 300
 DB 241 LGDRTGEVHHLYERDCSLQRRHQKVVETAPAOHLDPRLDRICADAVKFCRSIGYQAGT 300
 QY 301 VEFLVDEKGNHVFIEEMNPRIQVEHTVTEVTEVDLVKAQMRLAAGATLKLGLTQDKIKT 360
 DB 301 VEFLVDEKGNHVFIEEMNPRIQVEHTVTEVTEVDLVKAQMRLAAGATLKLGLTQDKIKT 360
 QY 361 HCAALQCRITTEDPNNGRPDGTITAYRSPGGAGVRLDGAAQLGGETTAHFDSMLVKMT 420
 DB 361 HCAALQCRITTEDPNNGRPDGTITAYRSPGGAGVRLDGAAQLGGETTAHFDSMLVKMT 420
 QY 421 CRGSDFEFAVARAALAEFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQA 480
 DB 421 CRGSDFEFAVARAALAEFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQA 480

QY 481 PPADDEQGRILDYLDADVTNPKPHGVPRKDVAAPIDKLPNIKDLPLPRGSRDLKOLGPAA 540
 DB 481 PPADDEQGRILDYLDADVTNPKPHGVPRKDVAAPIDKLPNIKDLPLPRGSRDLKOLGPAA 540
 QY 541 FARDLREODALAVTDTTFRDAHOSLLATRVRSFALKPAAEAVAKITPELLSVEANGGATY 600
 DB 541 FARDLREODALAVTDTTFRDAHOSLLATRVRSFALKPAAEAVAKITPELLSVEANGGATY 600
 QY 601 DYAMRFLFEDPWDRLEDELRAMPNNVNIOMLLRGRNTVGYTTPDSVCRAFVKEAASSVD 660
 DB 601 DYAMRFLFEDPWDRLEDELRAMPNNVNIOMLLRGRNTVGYTTPDSVCRAFVKEAASSVD 660
 QY 661 IFRIEDALNDVYSQMRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYLYLKMAEIV 720
 DB 661 IFRIEDALNDVYSQMRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYLYLKMAEIV 720
 QY 721 KSGAHILAIKDMAGLLRPAAVTKLTALRRREFDLRVHVTHTDTAGGOLATYFAAAQAGAD 780
 DB 721 KSGAHILAIKDMAGLLRPAAVTKLTALRRREFDLRVHVTHTDTAGGOLATYFAAAQAGAD 780
 QY 781 AVDGASAPLSGTTTSSPSLSAIVAAFAHTRRDTGLSLEAVSDELPYWEAVRGLYLPFESGT 840
 DB 781 AVDGASAPLSGTTTSSPSLSAIVAAFAHTRRDTGLSLEAVSDELPYWEAVRGLYLPFESGT 840
 QY 841 PGPTGRVYRHEIPGQSLNLRQAATLGLADRFELIEDNYAAVNEMLRGPKTKVTPSSKVV 900
 DB 841 PGPTGRVYRHEIPGQSLNLRQAATLGLADRFELIEDNYAAVNEMLRGPKTKVTPSSKVV 900
 QY 901 GDALHLVAGVDPADPADPOKYDIPDSVIAFLRGLGNPGGWPPEPLRTRALEGRSEG 960
 DB 901 GDALHLVAGVDPADPADPOKYDIPDSVIAFLRGLGNPGGWPPEPLRTRALEGRSEG 960
 QY 961 KAPLTVPEEEOAHLDDADDKSKRRNSLNRLFPKTESEFLEHRRRFGNTSALDDREFFYG 1020
 DB 961 KAPLTVPEEEOAHLDDADDKSKRRNSLNRLFPKTESEFLEHRRRFGNTSALDDREFFYG 1020
 QY 1021 LVEGRETLRLPDVPTPLLRDLDAISEPDDKCMRNVNANVNGOIRPMVRDRSVESVTAT 1080
 DB 1021 LVEGRETLRLPDVPTPLLRDLDAISEPDDKCMRNVNANVNGOIRPMVRDRSVESVTAT 1080
 QY 1081 AEKADSSNKGHVAAPFAGVYVTVTAEGDEVKAGDAVAIIIEAMKMEATITASVDGKIERVV 1140
 DB 1081 AEKADSSNKGHVAAPFAGVYVTVTAEGDEVKAGDAVAIIIEAMKMEATITASVDGKIERVV 1140
 QY 1141 VPAATKVEGGDLIVVVS 1157
 DB 1141 VPAATKVEGGDLIVVVS 1157
 RESULT 2
 AAU98052
 ID AAU98052 standard; Protein; 1140 AA.
 XX
 AC AAU98052;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 DE Corynebacterium mutant feedback-resistant pyruvate carboxylase enzyme.
 XX
 KW Feedback-resistant; pyruvate carboxylase; enzyme;
 KW aspartic acid feedback inhibition resistant; mutant; muten.
 XX
 OS Corynebacterium glutamicum.
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FT Misc-difference 1 /note= "Wild-type Met substituted by Val"
 FT Misc-difference 153 /note= "Wild-type Glu substituted by Asp"
 FT Misc-difference 182 /note= "Wild-type Ala substituted by Ser"
 FT

FT Misc-difference 206 /note= "Wild-type Ala substituted by Ser"
 FT Misc-difference 227 /note= "Wild-type His substituted by Arg"
 FT Misc-difference 455 /note= "Wild-type Ala substituted by Gly"
 FT Region 1110..1122 /note= "Specifically claimed in claim 18"
 FT Misc-difference 1116 /note= "Wild-type Asp substituted by Glu"
 XX
 PN WO200231158-A2.
 XX 18-APR-2002.
 XX
 XX 12-OCT-2001; 2001WO-US31893.
 XX
 XX 13-OCT-2000; 2000US-239913P.
 PR (ARCH) ARCHER-DANIELS MIDLAND CO.
 XX
 XX Hanke PD;
 XX
 DR WPI; 2002-463267/49.
 XX
 XX Novel mutated, feedback resistant pyruvate carboxylase enzyme
 PT polypeptide, useful for producing amino acids e.g. L-lysine,
 PT L-threonine, L-glycine, L-glutamic acid, L-proline and L-methionine and
 PT L-isoleucine
 XX
 PS Claim 1; Page -: 42pp; English.
 XX
 CC The present invention relates to a new mutated, feedback-resistant
 CC pyruvate carboxylase enzyme. The invention is useful for producing an
 CC amino acid (e.g. L-lys, L-thr, L-Met, L-Ile, L-Glu, L-Arg and L-Pro),
 CC by culturing a host cell in a suitable media and separating the amino
 CC acid from the medium. The vector of the invention is useful for
 CC replacement of a wild-type pyruvate carboxylase gene, with a feedback
 CC resistant pyruvate carboxylase gene, in a Corynebacterium glutamicum, by
 CC replacing a genomic copy of the wild-type pyruvate carboxylase gene with
 CC a selectable marker gene through homologous recombination to form a first
 CC recombination strain, and replacing the selectable marker gene in the
 CC first recombinant strain, with feedback resistant pyruvate carboxylase
 CC gene through homologous recombination to form a second recombinant
 CC strain, where the homologous recombination in the above steps, occurs
 CC between the host cell and the vector. The feedback-resistant pyruvate
 CC carboxylase enzyme is resistant to feedback inhibition from aspartic
 CC acid. The present amino acid sequence represents the mutant
 CC feedback-resistant pyruvate carboxylase enzyme of the invention.
 CC Note: The present sequence is not shown in the specification but is
 CC derived from the wild-type feedback-resistant pyruvate carboxylase
 CC enzyme (AAU98053) given in figure 2 of the specification.
 XX
 SQ Sequence 1140 AA;
 Query Match 98.5%; Score 5779; DB 23; Length 1140;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1138; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 18 VSTHTSSTLPAPKILVANRGEIAVRAALETGAATVAIYPREDRGSFHRSEFAEAVR 77
 DB 1 VSTHTSSTLPAPKILVANRGEIAVRAALETGAATVAIYPREDRGSFHRSEFAEAVR 60
 QY 78 ICTEGSPVKAYDIDIEIGAANKVKRADAIYPCYGFISENAQLARECAENGITFTGPTPEV 137
 DB 61 ICTEGSPVKAYDIDIEIGAANKVKRADAIYPCYGFISENAQLARECAENGITFTGPTPEV 120
 QY 138 LDLTGDKSRVTAANKAGLPVLAESTPSKNIDDIVKSAEGQYPIPVKAVAGGGGRGMRF 197
 DB 121 LDLTGDKSRVTAANKAGLPVLAESTPSKNIDDIVKSAEGQYPIPVKAVAGGGGRGMRF 180
 QY 198 VSSPDELKRLATEASREAEAFGDGSSVYVERAVINPQHIEVQILGDRTEGEVHLXERDCS 257
 DB 181 VSSPDELKRLATEASREAEAFGDGSSVYVERAVINPQHIEVQILGDRTEGEVHLXERDCS 240
 QY 258 LORRHOKVVEIAPAOHLDPDLRDRICADAVKFCRSIGYOGAGTVEFLVDEKGNHVEIEN 317
 DB 241 LORRHOKVVEIAPAOHLDPDLRDRICADAVKFCRSIGYOGAGTVEFLVDEKGNHVEIEN 300
 QY 318 PRIQVEHTVTEEVVDLVKAQMLAAGATLKEGLTQDKIKTHGAALOCRTTDPNNG 377
 DB 301 PRIQVEHTVTEEVVDLVKAQMLAAGATLKEGLTQDKIKTHGAALOCRTTDPNNG 360
 QY 378 FRPDTGTITAYRSPGGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDETAVARAQRAL 437
 DB 361 FRPDTGTITAYRSPGGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDETAVARAQRAL 420
 QY 438 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDPHLLQAPPADDDQGRITLDYADV 497
 DB 421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDPHLLQAPPADDDQGRITLDYADV 480
 QY 498 TVNKPCHGVPRKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAFAFARDLREQDALAVTDTT 557
 DB 481 TVNKPCHGVPRKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAFAFARDLREQDALAVTDTT 540
 QY 558 FRDAHQSLLATRVRSFALKPAEAAVAKLTPELLSVEAWGATYDVAMRFLFEDPMDRLDE 617
 DB 541 FRDAHQSLLATRVRSFALKPAEAAVAKLTPELLSVEAWGATYDVAMRFLFEDPMDRLDE 600
 QY 618 LREAMPNVNIOMLLRCRNTVGYTPYDPSVCRAFVKEAASGVDFIRIFDALNDVSOMRPA 677
 DB 601 LREAMPNVNIOMLLRCRNTVGYTPYDPSVCRAFVKEAASGVDFIRIFDALNDVSOMRPA 660
 QY 678 IDAVLETNTAVAEVAMAYSGDLPNEKLYTLDDYILKMAEIEVKSGAHLATIKDMAGLLR 737
 DB 661 IDAVLETNTAVAEVAMAYSGDLPNEKLYTLDDYILKMAEIEVKSGAHLATIKDMAGLLR 720
 QY 738 PAAVTKLVTLARREEDLPVHVHTHDTAGQLATYFAAAQAGADAVDGAAPLSGTTSQPS 797
 DB 721 PAAVTKLVTLARREEDLPVHVHTHDTAGQLATYFAAAQAGADAVDGAAPLSGTTSQPS 780
 QY 798 LSAIVAAFAHTRDTGLSLSEAVSDLEPYWEAVRGLVLPFESGTPGPTGRVYRHEIPGGOL 857
 DB 781 LSAIVAAFAHTRDTGLSLSEAVSDLEPYWEAVRGLVLPFESGTPGPTGRVYRHEIPGGOL 840
 QY 858 SNLRAQATGALGRFELIEDNYAAVNEMLGRPTKVTPTSSKVVGDALHLVAGVDPADF 917
 DB 841 SNLRAQATGALGRFELIEDNYAAVNEMLGRPTKVTPTSSKVVGDALHLVAGVDPADF 900
 QY 918 AADPKYDIPDSVIAFLRGLGNPPGWPPELTRALEGRSEKAPLTVPEEEQAHLDA 977
 DB 901 AADPKYDIPDSVIAFLRGLGNPPGWPPELTRALEGRSEKAPLTVPEEEQAHLDA 960
 QY 978 DSKERNLSNRLLPKPTTEEFLEHRRRFGNTSALDDREFFYGLVGEVRETILRLPDVRTP 1037
 DB 961 DSKERNLSNRLLPKPTTEEFLEHRRRFGNTSALDDREFFYGLVGEVRETILRLPDVRTP 1020
 QY 1038 LLVRLDAISEPDDKGMNRNVANVNGQIRPMVRDRSVESVTATAEKADSSNKGHVAAPFA 1097
 DB 1021 LLVRLDAISEPDDKGMNRNVANVNGQIRPMVRDRSVESVTATAEKADSSNKGHVAAPFA 1080
 QY 1098 GVTVTVTAEGDEVKAGDAVAIIEMKMEATITASVDCKITERVVVPAATKVEGDLIVVVS 1157
 DB 1081 GVTVTVTAEGDEVKAGDAVAIIEMKMEATITASVEGKIDRVVVVPAATKVEGDLIVVVS 1140
 RESULT 3
 AAB01436
 ID AAB01436 standard; Protein: 1141 AA.
 XX
 AC AAB01436;
 XX
 DT 20-OCT-2000 (first entry)
 XX
 DE Pyruvate carboxylase of C. glutamicum.
 XX

KW Pyruvate carboxylase; expression: amino acid biosynthesis; lysine;
KW glutamic acid; oxaloacetate; fermentation; biosynthesis.
XX Corynebacterium glutamicum.
PN WO200039305-A1.
XX 06-JUL-2000.
XX
XX
XX 23-DEC-1998; 98WO-US27301.
XX
XX 23-DEC-1998; 98WO-US27301.
XX
XX (SINS/) SINSKEY A J.
PA (LESS/) LESSARD P A.
PA (WILL/) WILLIS L B.
XX
PI Sinskey AJ, Lessard PA, Willis LB;
XX WPI: 2000-465746/40.
DR N-PSDB; AAA47533.
XX
XX Novel polynucleotides encoding Corynebacterium glutamicum pyruvate
PT carboxylase useful for industrial fermentation processes comprises a
PT specific nucleotide sequence
XX
XX Claim 3; Fig 1; 51pp; English.
XX
XX The pyruvate carboxylase of Corynebacterium glutamicum can be used
CC for producing amino acids, preferably lysine and glutamic acid in
CC industrial fermentations and for replenishing oxalacetate consumed
CC for biosynthesis during growth. By incorporating the pyruvate
CC carboxylase gene in expression vectors levels of expression can be
CC 2 - 20 fold higher than in Corynebacterium glutamicum.
XX
XX Sequence 1141 AA;
SQ

Query Match 98.3%; Score 5764; DB 21; Length 1141;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1134; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 17 LVSTHTSTLPAFKKILVANRGEIATVAFRAALETGAATVAIYPRDRGFSFHRSEAV 76
Db :|||||
Qy 1 MVSTHTSTLPAFKKILVANRGEIATVAFRAALETGAATVAIYPRDRGFSFHRSEAV 60
Db :|||||
Qy 77 RIGTEGSPVKAYLDIDEIGAANKVKADAIYPGVGLSENAQLARECAENGIITFIGTPE 136
Db :|||||
Qy 61 RIGTEGSPVKAYLDIDEIGAANKVKADAIYPGVGLSENAQLARECAENGIITFIGTPE 120
Db :|||||
Qy 137 VLDLTGDKSRVTAANKAGLPVLAESTPSKNIDIVKSAEGOTPIFVKAVAGGGRMR 196
Db :|||||
Qy 121 VLDLTGDKSRVTAANKAGLPVLAESTPSKNIDIVKSAEGOTPIFVKAVAGGGRMR 180
Db :|||||
Qy 197 FVSSPDELKLTASREAEAFDGGSVYVERAVINPQHIEVQILGDRTEGVVHLVERDC 256
Db :|||||
Qy 181 FVASPDELKLTASREAEAFDGGSVYVERAVINPQHIEVQILGDRTEGVVHLVERDC 240
Db :|||||
Qy 257 SLQRHOKVVEIAPAQHLDPQLDRICADAVKFCRSYGYGAGTVEFLVDEKGNHVFIE 316
Db :|||||
Qy 241 SLQRHOKVVEIAPAQHLDPQLDRICADAVKFCRSYGYGAGTVEFLVDEKGNHVFIE 300
Db :|||||
Qy 317 NPRIQVEHTVTEEVTEVDLVKAQMLAAGATLKEGLGTQDKIKTHGAALOCRIITDPNN 376
Db :|||||
Qy 301 NPRIQVEHTVTEEVTEVDLVKAQMLAAGATLKEGLGTQDKIKTHGAALOCRIITDPNN 360
Db :|||||
Qy 377 GFRPDTGITAYRSPGGVRLDGAALGGEITAHFDSMLVKMTCRGSDFFETAVARAQRA 436
Db :|||||
Qy 361 GFRPDTGITAYRSPGGVRLDGAALGGEITAHFDSMLVKMTCRGSDFFETAVARAQRA 420
Db :|||||
Qy 437 LAETVSGVATNIGFLRALLREEDFTSKRIATGFIQDHPHLLQAPPADDSQGRILDYLD 496
Db :|||||
Qy 421 LAETVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDSQGRILDYLD 480
Db :|||||

Qy 497 VTVNKPCHGVPRKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAAFARDLREODALAVTDT 556
Db :|||||
Qy 481 VTVNKPCHGVPRKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAAFARDLREODALAVTDT 540
Db :|||||
Qy 557 TFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGATYDVAMRFLFEDPWRLD 616
Db :|||||
Qy 541 TFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGATYDVAMRFLFEDPWRLD 600
Db :|||||
Qy 617 ELREAMPNVNIOMLLURGRNTVGYTPYDPSVCRFAFVKEAASGVDFIRIFDALNDSQMRP 676
Db :|||||
Qy 601 ELREAMPNVNIOMLLURGRNTVGYTPYDPSVCRFAFVKEAASGVDFIRIFDALNDSQMRP 660
Db :|||||
Qy 677 AIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYILKMAEIEIVKSGAHILAIDMAGLL 736
Db :|||||
Qy 661 AIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYILKMAEIEIVKSGAHILAIDMAGLL 720
Db :|||||
Qy 737 RPAAVTKLVTLRRREFDLPVHVHTHTAGGQATYFAAAQAGADAVDCASAPLSGTTTSP 796
Db :|||||
Qy 721 RPAAVTKLVTLRRREFDLPVHVHTHTAGGQATYFAAAQAGADAVDCASAPLSGTTTSP 780
Db :|||||
Qy 797 SLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGQ 856
Db :|||||
Qy 781 SLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGQ 840
Db :|||||
Qy 857 LSNLRAQATLGLADRFELIEDNVAAVNEMGLRPTKVTSSKVVGDALHLVGAGVDFAD 916
Db :|||||
Qy 841 LSNLRAQATLGLADRFELIEDNVAAVNEMGLRPTKVTSSKVVGDALHLVGAGVDFAD 900
Db :|||||
Qy 917 FAADPKYDIPDSVTAIFRGELGNPGGPPLRTRALEGRSECKAPLTVPEEQAHLD 976
Db :|||||
Qy 901 FANDPKYDIPDSVTAIFRGELGNPGGPPLRTRALEGRSECKAPLTVPEEQAHLD 960
Db :|||||
Qy 977 ADDSKERNNSNLLFPKPTTEFLHRRRFGNTSALDDREFFYGLVGEHRLIRLPDVRT 1036
Db :|||||
Qy 961 ADDSKERNNSNLLFPKPTTEFLHRRRFGNTSALDDREFFYGLVGEHRLIRLPDVRT 1020
Db :|||||
Qy 1037 PLLVRLDITSEPDCKGMNVVANVNGQIRPMRVDRDSVESVTATAEKADSSNKGHVAAPF 1096
Db :|||||
Qy 1021 PLLVRLDITSEPDCKGMNVVANVNGQIRPMRVDRDSVESVTATAEKADSSNKGHVAAPF 1080
Db :|||||
Qy 1097 AGVVTVTVAEGDEVKAGDAVAIIIEAMKMEATITASVDGKIERYVVVPAATKVEGGDLIVV 1156
Db :|||||
Qy 1081 AGVVTVTVAEGDEVKAGDAVAIIIEAMKMEATITASVDGKIERYVVVPAATKVEGGDLIVV 1140
Db :|||||
Qy 1157 S 1157
Db :
Qy 1141 S 1141
Db :
RESULT 4
AAG90511
ID AAG90511 standard; Protein; 1140 AA.
XX
XX AAG90511;
AC
XX
XX 26-SEP-2001 (first entry)
DT
XX
XX C glutamicum protein fragment SEQ ID NO: 4265.
DE
XX
XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis.
XX
XX Corynebacterium glutamicum.
OS
XX
XX EP1108790-A2.
PN
XX
XX 20-JUN-2001.
PD
XX
XX 18-DEC-2000; 2000EP-0127688.
PF
XX
XX 16-DEC-1999; 99JP-0377484.
PR
XX
XX 07-APR-2000; 2000JP-0159162.
PR
XX
XX 03-AUG-2000; 2000JP-0280988.
PR

XX (KYOW) KYOWA HAKKO KOGYO KK.
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A.
XX WPI: 2001-376931/40.
DR N-PSDB; AAF65730.
XX Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analysing
PT expression profile or pattern of a gene and identifying homologous gene
PT
XX Claim 17: SEQ ID NO: 4265; 246pp + Sequence Listing; English.
XX The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and
CC analysing the expression profile or expressing pattern of a gene derived
CC from Coryneform bacterium, and identifying a homologue of a gene derived
CC from coryneform bacterium. Coryneform bacteria are useful for producing
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described
CC in the exemplification of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
XX Sequence 1140 AA;
SQ Query Match 98.2%; Score 5759; DB 22; Length 1140;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1133; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
QY 18 VSTHTSSLTPAPKILVANRGEIAVRAALETGAATVAIYPREDRGSFHRFSASEAVR 77
DB 1 MTHSTSSLTPAPKILVANRGEIAVRAALETGAATVAIYPREDRGSFHRFSASEAVR 60
QY 78 ICTGSPKAYLIDIEITGAARKKADATYPCYGFLENAQALRECAENGITFIPTPEV 137
DB 61 ICTGSPKAYLIDIEITGAARKKADATYPCYGFLENAQALRECAENGITFIPTPEV 120
QY 138 LDLTGDKSRVTAAKKAGLPVLAESTPSKNIDDIIVKSAEGQYPIFVKAVAGGGGRGMRP 197
DB 121 LDLTGDKSRVTAAKKAGLPVLAESTPSKNIDDIIVKSAEGQYPIFVKAVAGGGGRGMRP 180
QY 198 VSSPDLKRLATEASREAEAFSGDGVYVERAVINPQHIEVQILGDRTEGVVHLYERDCS 257
DB 181 VASPDRLKRLATEASREAEAFSGDGVYVERAVINPQHIEVQILGDRTEGVVHLYERDCS 240
QY 258 LQRRHKQVVEIAPAOHLDELDRICADAVKFCRSIGYOGAGTVFELVDEKGNHVFIEHN 317
DB 241 LQRRHKQVVEIAPAOHLDELDRICADAVKFCRSIGYOGAGTVFELVDEKGNHVFIEHN 300
QY 318 PRIQVEHTVTEEVTEVDLVKAQMLAAGATLKELGITQDKIKTHGAALQCRITTEDPNNG 377
DB 301 PRIQVEHTVTEEVTEVDLVKAQMLAAGATLKELGITQDKIKTHGAALQCRITTEDPNNG 360
QY 378 FRPDTGTTAYRSPGAGVRLDGAQOLGGEITAHFDSMLVKMTCRGSDFETAVARAQAL 437
DB 361 FRPDTGTTAYRSPGAGVRLDGAQOLGGEITAHFDSMLVKMTCRGSDFETAVARAQAL 420
QY 438 AEFTVSGVATNIGFURALLREEDFTSKRIATGFIQDHPHLLQAPPADDEQGRILDYLDV 497
DB 421 AEFTVSGVATNIGFURALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLDV 480
QY 498 TVNKHGVRPKDVAAPIDKLPNIKDLPLRSGSRDLKQLGPAFAFADLREQDALAVTDTT 557
DB 481 TVNKHGVRPKDVAAPIDKLPNIKDLPLRSGSRDLKQLGPAFAFADLREQDALAVTDTT 540
QY 558 FRDAHOSLLATRVRSFALKPAAEAVAKLTPELLSVSEAMGGATYDVAMRFLFEDPWRDLDE 617

DB 541 FRDAHOSLLATRVRSFALKPAAEAVAKLTPELLSVSEAMGGATYDVAMRFLFEDPWRDLDE 600
QY 618 LREAMPNVNIQMLLRCGRNTVGYTPYDPSVCRAPFVKEAASGVYDIFRIFDALNDVSMRPA 677
DB 601 LREAMPNVNIQMLLRCGRNTVGYTPYDPSVCRAPFVKEAASGVYDIFRIFDALNDVSMRPA 660
QY 678 IDAVLETTNTAVAEVAMAYSGDLSDPNEKLYTLDYILKMAEIEVKSGAHLATKDMAGLLR 737
DB 661 IDAVLETTNTAVAEVAMAYSGDLSDPNEKLYTLDYILKMAEIEVKSGAHLATKDMAGLLR 720
QY 738 PAAVTKLYLTALRREFDLPVHVHTHDTAGQLATYFAAAQAGADAVDGSAPLSGTTSQPS 797
DB 721 PAAVTKLYLTALRREFDLPVHVHTHDTAGQLATYFAAAQAGADAVDGSAPLSGTTSQPS 780
QY 798 LSAIAVAFAHTRDTCLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYHIEIPGGOL 857
DB 781 LSAIAVAFAHTRDTCLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYHIEIPGGOL 840
QY 858 SNLRAQATGALGLADREFELIEDNYAAVNEMLGRPTKVTSSKVVGDIALHLVAGVDPADF 917
DB 841 SNLRAQATGALGLADREFELIEDNYAAVNEMLGRPTKVTSSKVVGDIALHLVAGVDPADF 900
QY 918 AADPOKYDIPDSVIAFLRQELGNPPGGWPEPLRTRALEGRSEKAPLTVPEEEOAHLDA 977
DB 901 AADPOKYDIPDSVIAFLRQELGNPPGGWPEPLRTRALEGRSEKAPLTVPEEEOAHLDA 960
QY 978 DSKERRNSLNRLLPKPTTEEFLEHRRRFGNTSALDDREFFYGLVGEETLRLPDVPTP 1037
DB 961 DSKERRNSLNRLLPKPTTEEFLEHRRRFGNTSALDDREFFYGLVGEETLRLPDVPTP 1020
QY 1038 LLVRLDAISEPDDKGMNVYVANVNGQIRPMRVDRSVESVTATAEKADSSNKGHVAAPFA 1097
DB 1021 LLVRLDAISEPDDKGMNVYVANVNGQIRPMRVDRSVESVTATAEKADSSNKGHVAAPFA 1080
QY 1098 GVVTVTAEGDEVKAGDAVAIIEMKMEATITASVDGKIETRVVVPVPAATKVEGDLIVVYS 1157
DB 1081 GVVTVTAEGDEVKAGDAVAIIEMKMEATITASVDGKIETRVVVPVPAATKVEGDLIVVYS 1140
RESULT 5
AAB67129
ID AAB67129 standard; Protein; 1140 AA.
XX
AC AAB67129;
XX
DT 12-APR-2001 (first entry)
XX
DE Corynebacterium glutamicum pyruvate carboxylase.
XX
KW Pyruvate carboxylase; anaplerotic pathway; industrial fermentation;
KW oxaloacetate.
XX
OS Corynebacterium glutamicum.
XX
PN US6171833-B1.
PD 09-JAN-2001.
XX
PF 23-DEC-1998; 98US-0220081.
XX
PR 23-DEC-1998; 98US-0220081.
XX
PA (MASI) MASSACHUSETTS INST TECHNOLOGY.
XX
PI Sinskey AJ, Lessard PA, Willis LB;
XX
DR WPI: 2001-122330/13.
XX N-PSDB; AAF32165.
PT Novel nucleic acid encoding pyruvate carboxylase from Corynebacterium
PT glutamicum, for replenishing oxaloacetate consumed during lysine and
PT glutamic acid production in industrial fermentations -

QY	1038	LLVRLDAISEPDDKGRNVVAVNGVQIRPMVRDRSVESVTATAEKADSSNKGHVAAPFA	1097
Db	1021	LLVRLDAISEPDDKGRNVVAVNGVQIRPMVRDRSVESVTATAEKADSSNKGHVAAPFA	1080
QY	1098	GVVTVTVABGVKAGDAVAIIAEAMKMEATITASVDGKIERVVVPAAATKVEGGDLIVVVS	1157
Db	1081	GVVTVTVABGVKAGDAVAIIAEAMKMEATITASVDGKIDRVVVPAAATKVEGGDLIVVVS	1140
RESULT 7			
AAW93971		AAW93971 standard; Protein; 1140 AA.	
XX		AC	
XX		AAW93971;	
DT	30-JUN-1999	(first entry)	
XX			
DE		C. glutamicum pyruvate carboxylase protein.	
XX			
KW		Pyruvate carboxylase; amino acid production; lysine production;	
KW		threonine production; homoserine production; glutamate production;	
KW		arginine production; feed additive; condiment; pharmaceutical;	
KW		fine chemical; ss.	
XX			
OS		Corynebacterium glutamicum.	
XX			
PN		DE19831609-A1.	
XX			
PD		15-APR-1999.	
XX			
PF		14-JUL-1998; 98DE-1031609.	
XX			
PR		04-OCT-1997; 97DE-1043894.	
XX			
PA		(KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.	
XX			
PI		Eikmanns B, Peters-Wendisch P, Sahm H;	
XX			
DR		WPI; 1999-245521/21.	
DR		N-PSDB; AAX24102.	
XX			
PT		Increasing microbial production of specific amino acids by	
PT		increasing activity or expression of pyruvate carboxylase	
XX			
PS		Claim 15; Page 11-15; 18pp; German.	
XX			
CC		This invention describes the isolation of a pyruvate carboxylase	
CC		from Corynebacterium glutamicum which is used in a novel method for	
CC		production of lysine, threonine, homoserine, glutamate and/or arginine	
CC		variously useful as feed additives, condiments, pharmaceuticals and	
CC		intermediates for fine chemicals. Increasing pyruvate carboxylase	
CC		activity increases the yield of microbial production of amino acids	
CC		of the aspartate and/or glutamate families, e.g. about 50% more lysine,	
CC		40% more threonine and 150% more homoserine are secreted into the	
CC		culture medium.	
XX			
SQ		Sequence 1140 AA;	
Query Match		98.1%; Score 5753; DB 20; Length 1140;	
Best Local Similarity		99.3%; Pred. No. 0;	
Matches 1132; Conservative		5; Mismatches 3; Indels 0; Gaps	
QY	18	VSTHTSSTLPAPFKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRFSAEAVR	77
Db	1	MGTHTSSTLPAPFKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRFSAEAVR	60
QY	78	IGTEGSPVKAYLIDIEIIIGAARKVKADAIYPGYFLSENAQLARECAENGITFIGPTPEV	137
Db	61	IGTEGSPVKAYLIDIEIIIGAARKVKADAIYPGYFLSENAQLARECAENGITFIGPTPEV	120
QY	138	LDLTGDKSRVTAARKAGLPVLAESTPSKNIDDIVKSAEGQTYPIFVKAVAGGGGRMRF	197

Db 121 LDLTGDKSRVATAAKKAGLPVLAESTPSKNIIDEIVKSAEGQTYPIFVKAVAGGGGRMRF 180
 Qy 198 VSSPDELRLKRLATEASREAEAFDGSVYVERAVINPOHIEVQILGDRTEGVVHLYERDCS 257
 Db 181 VASPDRLKRLATEASREAEAFDGSVYVERAVINPOHIEVQILGDRTEGVVHLYERDCS 240
 Qy 258 LORRHQKVETAPQAHLDPQLDRICADAVKFCRSIGYCGAGTVEFLVDEKGNHVEIWN 317
 Db 241 LORRHQKVETAPQAHLDPQLDRICADAVKFCRSIGYCGAGTVEFLVDEKGNHVEIWN 300
 Qy 318 PRIOVEHTVTEVTEVDLVKRAOMRLAAGATLKLGLTQDKIKTHGAALQCRITTEPNNG 377
 Db 301 PRIOVEHTVTEVTEVDLVKRAOMRLAAGATLKLGLTQDKIKTHGAALQCRITTEPNNG 360
 Qy 378 PRDGTGITAYRSPGGAGVRLDGAQJGGEITAHFOSMLVKMTCRGSDFETAVARAQAL 437
 Db 361 PRDGTGITAYRSPGGAGVRLDGAQJGGEITAHFOSMLVKMTCRGSDFETAVARAQAL 420
 Qy 438 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIQDRPHLLQAPPADDEQGRILDYADV 497
 Db 421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIQDRPHLLQAPPADDEQGRILDYADV 480
 Qy 498 TVNPKHGVRRPKDVAAPIDKLPNFKDLPLPRGSRDLKQLGPAFAFARDLREQDALAVTDTT 557
 Db 481 TVNPKHGVRRPKDVAAPIDKLPNFKDLPLPRGSRDLKQLGPAFAFARDLREQDALAVTDTT 540
 Qy 558 FRDAHQSILLATRVRSFALKPAEAAVAKLTPELLSVEAWGATYDVAMRFLFEDPWRDLDE 617
 Db 541 FRDAHQSILLATRVRSFALKPAEAAVAKLTPELLSVEAWGATYDVAMRFLFEDPWRDLDE 600
 Qy 618 LREAMPNVNTOMLRGRNTVGYTPYDPSVCRAFVKEAASGVDFIRFDALNDVDSOMRPA 677
 Db 601 LREAMPNVNTOMLRGRNTVGYTPYDPSVCRAFVKEAASGVDFIRFDALNDVDSOMRPA 660
 Qy 678 IDAVLENTVAEYAMAYSGDLSPNEKLYLDYLYKMAEEIVKSGAHLIAIKDMAGLLR 737
 Db 661 IDAVLENTVAEYAMAYSGDLSPNEKLYLDYLYKMAEEIVKSGAHLIAIKDMAGLLR 720
 Qy 738 PAAVTKLVTLRREFDLPVHVHHTDAGQOLATYFAAAQAGADAVDGSAPLSGTTSQPS 797
 Db 721 PAAVTKLVTLRREFDLPVHVHHTDAGQOLATYFAAAQAGADAVDGSAPLSGTTSQPS 780
 Qy 798 LSAIVAAFAHTRRTDGLSLLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGQOL 857
 Db 781 LSAIVAAFAHTRRTDGLSLLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGQOL 840
 Qy 858 SNLRAQATAGLADRFELIEDNVAANVEMLGRPTKVTTPSSKVYVGDALHLVAGVDPADF 917
 Db 841 SNLRAQATAGLADRFELIEDNVAANVEMLGRPTKVTTPSSKVYVGDALHLVAGVDPADF 900
 Qy 918 AADPKYDIPDSVIAFLRGLNPPGWPPEPLRTRALEGRSECKAPLTVPEEEQAHLDA 977
 Db 901 AADPKYDIPDSVIAFLRGLNPPGWPPEPLRTRALEGRSECKAPLTVPEEEQAHLDA 960
 Qy 978 DSKERNLSNRLLPKPTBEFFLEHRRFFGNTSALDDREFFYGLVGRFTLRLPDVPTP 1037
 Db 961 DSKERNLSNRLLPKPTBEFFLEHRRFFGNTSALDDREFFYGLVGRFTLRLPDVPTP 1020
 Qy 1038 LLVBLDAISBDDKGMNVNANVNGQIRPMVRDRSVESVTATAEKADSNKGHVAAPA 1097
 Db 1021 LLVBLDAISBDDKGMNVNANVNGQIRPMVRDRSVESVTATAEKADSNKGHVAAPA 1080
 Qy 1098 GVVTVTVAEGDEVKAGDAVALIEMKMEATITASVDGKIERRVVVPAATKVEGGDLIVVVS 1157
 Db 1081 GVVTVTVAEGDEVKAGDAVALIEMKMEATITASVDGKIERRVVVPAATKVEGGDLIVVVS 1140

RESULT 8
 ID AAG93249 standard; Protein: 1140 AA.
 XX
 AC AAG93249;
 XX

DT 26-SEP-2001 (first entry)
 XX C glutamicum protein fragment mutant P458S.
 DE
 XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
 KW organic acid synthesis; mutant; muten.
 XX Corynebacterium glutamicum.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 458 /note= "wild-type Pro substituted by Ser"
 FT
 XX
 PN EPI108790-A2.
 XX
 XX 20-JUN-2001.
 PD
 XX 18-DEC-2000; 2000EP-0127688.
 PF
 XX 16-DEC-1999; 99JP-0377484.
 PR 07-APR-2000; 2000JP-0159162.
 PR 03-AUG-2000; 2000JP-0280988.
 XX
 XX (KYOW) KYOWA HAKKO KOGYO KK.
 PA
 XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
 XX
 XX WPI; 2001-376931/40.
 DR
 XX Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analysing
 PT expression profile or pattern of a gene and identifying homologous gene
 PT
 XX
 PS Claim 43; Page -: 246pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and
 CC analysing the expression profile or expression pattern of a gene derived
 CC from Coryneform bacterium, and identifying a homologue of a gene derived
 CC from coryneform bacterium. Coryneform bacteria are useful for producing
 CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a mutant protein described
 CC in the exemplification of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC European Patent Office.
 XX
 SQ Sequence 1140 AA;
 Query Match 98.1%; Score 5751; DB 22; Length 1140;
 Best Local Similarity 99.3%; Pred. No. 0;
 Matches 1132; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
 Qy 18 VSTHTSSLPAPFKKLLVANRGEIARAFRAALTAETGAATVAIYPREDRGSFHRSEAVR 77
 Db 1 MSHHTSSLPAPFKKLLVANRGEIARAFRAALTAETGAATVAIYPREDRGSFHRSEAVR 60
 Qy 78 IGTGSPVKAYLIDDEIIGAARKKADAIYPGYFLSNAQLARECAENGITFTGPTPEV 137
 Db 61 IGTGSPVKAYLIDDEIIGAARKKADAIYPGYFLSNAQLARECAENGITFTGPTPEV 120
 Qy 138 LDLTGDKSRVATAAKKAGLPVLAESTPSKNIIDEIVKSAEGQTYPIFVKAVAGGGGRMRF 197
 Db 121 LDLTGDKSRVATAAKKAGLPVLAESTPSKNIIDEIVKSAEGQTYPIFVKAVAGGGGRMRF 180
 Qy 198 VSSPDELRLKRLATEASREAEAFDGSVYVERAVINPOHIEVQILGDRTEGVVHLYERDCS 257
 Db 181 VASPDRLKRLATEASREAEAFDGSVYVERAVINPOHIEVQILGDRTEGVVHLYERDCS 240


```
Db 421 AEFNSGVATNIGFLRALLREEDFTKRRIDTGFISGHQHLQAPPADDEQGRILEYLADV 480
QY 498 TVNKPCHVRPKQVAAPIDKLPNIDKLPGRSGRDLKQLGPAAFARDLREQDALAVDTDT 557
Db 481 TVNKPCHGERP-ETARPIEKLPVEINPLPRGSRDLKQLGPGEFARDLREQDALAVDTDT 539
QY 558 FRAHQSLATRVRSFALPAAEAVAKLTPELLSSVEAWGATYDVAMRFLFEDPMDRLDE 617
Db 540 FRAHQSLATRVRSFALTPAARAVAKLTPELLSSVEAWGATYDVAMRFLFEDPMDRLDE 599
QY 618 LREAMPNVNLOMLLRGNTVGYTPYDSCVCRAPVKEAASGVVDIFRFDALNDVSMRPA 677
Db 600 LREAMPNVNLOMLLRGNTVGYTPYDSCVCRAPVQRAASGVVDIFRFDALNDVSMRPA 659
QY 678 IDAVLETNTAAVAVAMAYSGDLSNPKEKLYTLDYILKMAEIVKSGAHLAIKDMAGLLR 737
Db 660 IDAVLETNTAAVAVAMAYSGDLSNPKEKLYTLDYILKMAEIVKSGAHLAIKDMAGLLR 719
QY 738 PAAVTKLVTLARREFDLPVHVHTHDTAGGOLATYFAAAGAGADAVDGAAPLSGTTSPQS 797
Db 720 RAAAPKLVTALREFDLPVHVHTHDTAGGOLATYFAAAGAGADAVDGAAPLSGTTSPQS 779
QY 798 LSAIAFAHTRDRTGLSLAEVSDLEPYWEAVRGLYLPFESGTPGTRVYRHEIPGGOL 857
Db 780 MSALVAFAHTRDRTGLNQAQVSDLEPYWEAVRGLYLPFESGTPGTRVYRHEIPGGOL 839
QY 858 SNLRAQATLGLADREFELIEDNTAAVNEMLGRPTKVTSSKVVGDIALHLVAGVDPADF 917
Db 840 SNLRAQAVLGLADREFELIEDNTAAVNEMLGRPTKVTSSKVVGDIALHLVAGVSPDEF 899
QY 918 AADPKQYDIPDSVIAFLRGLNPPGWPPEPLTRALEGRSGKAPLTVPEEQAHLDA 977
Db 900 AADPKQYDIPDSVIAFLRGLNPPGWPPEPLTRALEGRSGKAPLTVPEEQAHLDS 959
QY 978 DSKERRNSLNRLLFPKPTPEEFLEHRRRGNTSALDDREFFYGLVGRETLRLPDVTRP 1037
Db 960 DSAERRGTLNRLFPKPTPEEFLEHRRRGNTSALDDREFFYGLVGRETLRLGVSTP 1019
QY 1038 LLYRLDAISPDDKGRNVVYVNGOIRPMRVDRSVESVTATAEKADSSNGHVAAPFA 1097
Db 1020 MVLRLDAVSEDDKGRNVVYVNGOIRPMRVDRSVESVTATAEKADATNGHVAAPFA 1079
QY 1098 GVVTVTVAEDEKAGDAVAIEMKMEATITASVDGKIVRVVPAATKVEGGDLIVVVS 1157
Db 1080 GVVTVTVAEDEKAGDAVAIEMKMEATITAPVDGVIDRVVVVPAATKVEGGDLIVVVS 1139

RESULT 10
AAB79302
ID AAB79302 standard; Protein; 532 AA.
XX
AC AAB79302;
XX
DT 30-APR-2001 (first entry)
XX
DE Corynebacterium glutamicum SMP protein sequence SEQ ID NO:120.
XX
KW Corynebacterium glutamicum; carbon metabolism and energy production;
KW SMP protein; sugar metabolism and oxidative phosphorylation protein;
KW fine chemical production; organic acid; proteinogenic amino acid;
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
KW carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;
KW diagnosis; Corynebacterium diphtheriae; evolutionary study.
XX
OS Corynebacterium glutamicum.
XX
PN WC200100844-A2.
XX
PD 04-JAN-2001.
XX
PF 23-JUN-2000; 2000MO-IB00943.
```

```
XX 25-JUN-1999; 99US-0141031.
PR 08-JUL-1999; 99DE-1031412.
PR 08-JUL-1999; 99DE-1031413.
PR 08-JUL-1999; 99DE-1031419.
PR 08-JUL-1999; 99DE-1031420.
PR 08-JUL-1999; 99DE-1031424.
PR 08-JUL-1999; 99DE-1031428.
PR 08-JUL-1999; 99DE-1031431.
PR 08-JUL-1999; 99DE-1031433.
PR 08-JUL-1999; 99DE-1031434.
PR 08-JUL-1999; 99DE-1031510.
PR 08-JUL-1999; 99DE-1031562.
PR 08-JUL-1999; 99DE-1031634.
PR 09-JUL-1999; 99DE-1032180.
PR 09-JUL-1999; 99DE-1032227.
PR 09-JUL-1999; 99DE-1032230.
PR 09-JUL-1999; 99US-0143208.
PR 14-JUL-1999; 99DE-1032924.
PR 14-JUL-1999; 99DE-1032973.
PR 14-JUL-1999; 99DE-1033005.
PR 27-AUG-1999; 99DE-1040765.
PR 31-AUG-1999; 99US-0151572.
PR 03-SEP-1999; 99DE-1042076.
PR 03-SEP-1999; 99DE-1042079.
PR 03-SEP-1999; 99DE-1042086.
PR 03-SEP-1999; 99DE-1042087.
PR 03-SEP-1999; 99DE-1042088.
PR 03-SEP-1999; 99DE-1042095.
PR 03-SEP-1999; 99DE-1042123.
PR 03-SEP-1999; 99DE-1042125.
XX
XX (BADI ) BASF AG.
XX
PI Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
XX
XX WPI: 2001-061975/07.
XX N-PSDB; AAF71419.
XX
XX New isolated Corynebacterium glutamicum nucleic acid encoding a sugar
XX metabolism and oxidative phosphorylation protein for production or
XX modulation of production of fine chemicals e.g. amino acids,
XX carbohydrates or enzymes
XX
XX Claim 20; Page 319-320; 1246pp; English.
XX
XX AAF71360 to AAF71750 encode the Corynebacterium glutamicum sugar
XX metabolism and oxidative phosphorylation (SMP) proteins given in
XX AAB79243 to AAB 79633 which are involved in carbon metabolism and
XX energy production. The C. glutamicum SMP gene can be used in vectors
XX (II) for expression in host cells and production or modulation of
XX production of fine chemicals, such as, an organic acid, a proteinogenic
XX or nonproteinogenic amino acid (preferred), a purine or pyrimidine base,
XX a nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty
XX acid, a diol, a carbohydrate, an aromatic compound, a vitamin, a
XX cofactor, a polyketide, or an enzyme. The presence of (I) or SMP proteins
XX (III) encoded by them are used for diagnosing the presence or activity of
XX Corynebacterium diphtheriae in a subject. (I), (II), (III) or host cells
XX containing them are used to map genomes of organisms related to
XX C. glutamicum, identify and localise C. glutamicum sequences of interest,
XX in evolutionary studies, in determining SMP protein regions required
XX for function, in modulating SMP protein activity in modulating the
XX metabolism of sugars, and in modulating high-energy molecule production
XX in a cell (i.e. ATP, NADPH).
XX
XX Sequence 532 AA;
XX
XX Query Match 45.7%; Score 2678; DB 22; Length 532;
XX Best Local Similarity 98.9%; Pred. No. 6.6e-191;
XX Matches 526; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
XX
XX 1 MTAITGLGLLKGIITLVSTHTSSTLPAFKKILVANKEIAVRAFRALLETGAATVAIYP 60
XX :|||||
XX
```

Db 1 VTAITLGGILLKGIITLVSTHTSSTLPAPFKILVANRGEIAVRAFRALLETGAATVAIYP 60
QY 61 REDRGSFHRSEAFSEAVRIGTEGSPVKAYLIDIEIIIGAAKKVADAIYPGYFLSENAQLA 120
Db 61 REDRGSFHRSEAFSEAVRIGTEGSPVKAYLIDIEIIIGAAKKVADAIYPGYFLSENAQLA 120
QY 121 RECAENGITFIGTPTPEVLDTGDKSAVTAAKKAGLPVLAESTPSSKNIDDIIVKSAEGQTY 180
Db 121 RECAENGITFIGTPTPEVLDTGDKSAVTAAKKAGLPVLAESTPSSKNIDDIIVKSAEGQTY 180
QY 181 PIFVKAVAGGGGMRFEVSSPDELRLKATEASREAEAAFGDGSVYVERAVINPOHIEVOI 240
Db 181 PIFVKAVAGGGGMRFEVSSPDELRLKATEASREAEAAFGDGSVYVERAVINPOHIEVOI 240
QY 241 LGDRTGEVHLVERDCSLORRHQKVVEIAPAOHLDPDELDRICADAVKFCRSIGYGAGT 300
Db 241 LGDHTGEVHLVERDCSLORRHQKVVEIAPAOHLDPDELDRICADAVKFCRSIGYGAGT 300
QY 301 VEFVLDEKGNHVFIEIENPRIQVEHTVEEVTEVDLVKAQMRLAAGATLKGELGLTQDKIKT 360
Db 301 VEFVLDEKGNHVFIEIENPRIQVEHTVEEVTEVDLVKAQMRLAAGATLKGELGLTQDKIKT 360
QY 361 HGAALQCRITTEDPNNGFPPDGTITAYRSPGGAGVRLDGAALGGEITAHFDSMLVKMT 420
Db 361 HGAALQCRITTEDPNNGFPPDGTITAYRSPGGAGVRLDGAALGGEITAHFDSMLVKMT 420
QY 421 CRGSDPETAARAQALAEFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPLLQA 480
Db 421 CRGSDPETAARAQALAEFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPLLQA 480
QY 481 PPADDEQGRILDYLAADVTVNKPGRPKDVAAPIDKLPNIKDLPLPRGSRDR 532
Db 481 PPADDEQGRILDYLAADVTVNKPGRPKDVAAPIDKLPNIKDLPLPRGSRDR 532
RESULT 11
AAU00511
ID AAU00511 standard; Protein; 1148 AA.
XX AAU00511;
AC AAU00511;
XX 07-SEP-2001 (first entry)
XX Bacillus subtilis pyruvate carboxylase enzyme A.
XX Pyruvate carboxylase A; pyCA; oxaloacetate; Escherichia;
KW L-amino acid production; fermentation.
XX Bacillus subtilis strain 168.
XX EP1092776-A1.
XX 18-APR-2001.
XX 05-OCT-2000; 2000EP-0121763.
XX 14-OCT-1999; 99RU-0121636.
XX (AJIN) AJINOMOTO CO INC.
XX Gussyatiner MM, Kozlov YI, Ptityn LR, Altman IB, Voroshilova EB;
PI Iomantas YAV, Yampolskaya TA;
XX WPI: 2001-309819/33.
DR N-PSDB; AAS01509.
XX New bacterium from the genus Escherichia containing a gene encoding for
PT pyruvate carboxylase useful for producing higher concentrations of
PT L-amino acids
XX Claim 4; Page 17-21; 28pp; English.
PS The present sequence represents Bacillus subtilis pyruvate
CC

CC carboxylase enzyme A. The pyCA gene encodes for pyruvate carboxylase
CC which catalyses the carboxylation of pyruvate to form oxaloacetate.
CC Transformation of the Bacillus subtilis pyCA gene into a bacterium
CC belonging to the genus Escherichia results in the bacterium showing
CC L-amino acid productivity. The invention provides a novel method for
CC producing an L-amino acid by fermentation. The method involves
CC cultivating the bacterium in a medium and producing and accumulating
CC the L-amino acid in the medium. The new bacterium harbouring the gene
CC coding for an enzyme having pyruvate carboxylase activity is useful
CC for producing higher concentrations of L-amino acids in vitro than
CC prior art.
XX Sequence 1148 AA;
SQ Query Match 43.1%; Score 2530; DB 22; Length 1148;
Best Local Similarity 46.0%; Pred. No. 2.2e-179;
Matches 533; Conservative 188; Mismatches 390; Indels 48; Gaps 13;
QY 28 AFKKILVANRGETAVRAFRALLETGAATVAIYPREDRGSFHRSEAFSEAVRIGTEGSPVKA 87
Db 5 SIQKVLVANRGETAIRIFRACTELNIRTVAVYSKEDSGSYHRYKADEAVLGVGKKPIDA 64
QY 88 YLDIDIEIIIGAAKKVADAIYPGYFLSENAQLARECAENGITFIGTPTPEVLDTGDKSRA 147
Db 65 YLDIEGIIIDIAKRNVDAIHPGYFLSENIHFARRCEEGIVFIGPKSEHLMFGDKVKA 124
QY 148 VTAACKAGLPVLAEST-PSKNIDDIIVKSAEGQTYPIFVKAVAGGGGMRFEVSSPDELRLK 206
Db 125 RQAEKAGIPVPGSDGPAETLEAVEQFGAANGYPIIIIRASLGGGGGMRIVRSESEVKE 184
QY 207 LATEASREAEAAFGDGSVYVERAVINPOHIEVOILGDRTEGVVHLVERDCSLORRHQKV 266
Db 185 AYERAKSEAKAAGFNDDEVYVEKLIENPKHIEVQVIGDKQGNVHLFERDCSVQRHOKVI 244
QY 267 ETAPAOHLDPDELDRICADAVKFCRSIGYGAGTVEFLVDEKGNHVFIEIENPRIQVEHTV 326
Db 245 EVAPSVLSPELRDQICEAAALAKNVNINAGTVEFLV-ANNEFYFIEVNPVRVQVEHTI 303
QY 327 TEEVTEVDLVKAQMRLAAGATL--KELGLTQDK--IKTHGAALQCRITTEDPNNGFPPDGTG 383
Db 304 TEMITGVDIVQTIILVAQGHLSHKVNIPKQIDITIGYAIQSRVTTEDPQNDPMPDTG 363
QY 384 TITAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDPETAARAQALAEFTV 442
Db 364 KIMAYRSGGGFGVRLDTGNSFGAVITPYDLSLLYKLTWALTFFEQAAKKMYRNLOEFRI 423
QY 443 SGVATNIGFLRALLREEDFTSKRIATGFIGDHPLLQA PADDEQGRILDYLAADVTVN-- 500
Db 424 RGIKTNIPELVNAKHEKELTQGYDTSFIDTTFELNFPFKQKDRGTMKLTYYIGNVTYNGF 483
QY 501 -----KPHGVPRPKDVAAPIDKLPNIKDLPLPRGSRDRILKOLGPAFAAFARDLREQDALAV 553
Db 484 PGIGKKEKPAFKPLGVKVDVDDQP-----ARGTKQILDEKGAEGLANWKEQKSVLL 536
QY 554 TDTTFRDAHQSLLATRVRSFALKPAEAVAKLTPELLSVSEAMGGATYDVAMRFLFEDPDND 613
Db 537 TDTTFRDAHQSLLATRI RSHDLKKITANPTAALWPELFSMEMGGATFDVAYRFLKEDPNK 596
QY 614 RLDELREAMPNVNIQMLLRGNTVGYTPYDSCVCFVKEAAASSGVVDIFRFDALNDVSO 673
Db 597 RLEDLRKEVNTLTFQMLLRSSNAVGYTNPVNIKEFVAKQSAQSGIDVFRIEDSLANWVG 656
QY 674 MRPAIDAVLETNTVAEAVAMAYSGDLSDPNEKLYTLDDYLLKWAEEIVKSGAHTLAKDMA 733
Db 657 MTLAIDAVRDTG-KVAEAAICYTGDILDKNRTKYDLAYTSMAKELEAAAGAHILGLKDMA 715
QY 734 GLLRPAAVTKLVTALRREFDLPVHVHTHTAGGQATYFAAAQAGADADVASAPLGGTT 793
Db 716 GLLKQAAVELVSALKETIDIPVHLHTDTSNGITMYAKAVEAGVDIIDVAVSSMAGLT 775
QY 794 SQPSLSAIVAAFAHTRRDTGLSLEAVSDILEPYWEAVRGLYLPFESGTPGTGRVYRHEIP 853
Db 776 SQPSASGFYHAMEGNDRRPEMNVQGVELLQYVESVRYKYYSEFESGMSKSPHTEIYEHMP 835

QY 854 GGQSLNRAQATLGLADRELIEDNYAANVEMGLRPTKVTTPSSKVVGDIALHLVAGVD 913
 Db 836 GGQSLNRAQATLGLADRELIEDNYAANVEMGLRPTKVTTPSSKVVGDIALHLVAGVD 913
 QY 914 PADFAADQKYDIPDSVIAFLRGLGNPPGWPPEPLRTRALEGRSGKAPLTVPEE--- 970
 Db 896 EKDYKESGLDFPDSVVEFLKGNIGOPHGGFPEKQKLIKQOE---PITVRPELLE 951
 QY 971 -----EQALHDADDKSRNLSNRLFLFPKPTPEFLEHRRFGNTSALDDREF 1018
 Db 952 PVSPEAKIQEKEHNEISD---QDAVAYALYPKVFTDYVKTTSYGDLSLVDPTTF 1007
 QY 1019 YGLVEGRETILRLPDVRLPLRLDLSPEDDKGMNVMYANVQIRPMRDRSVESVT 1078
 Db 1008 YGMTLGEIEVEIERGT-LIVKLISIGEPQDPATRVVYFELNGQPREVVKDESISV 1066
 QY 1079 ATAEKADSNKGVHAPFAGVTVTVAE-GDEVKAGDAVAITAEKMEATITASVDGKIE 1137
 Db 1067 OERLKADRTNPSHIAASMPGTVIKVLAEAGTKVNGDHLINAMKMETVQAPFSGTIR 1126
 QY 1138 RVVVPAAKVEGGDLVVV 1156
 Db 1127 QVHVKNCEPIQTGDLLEI 1145

RESULT 12
 AB47612
 ID AB47612 standard; Protein: 1146 AA.
 AC AB47612;
 XX 05-FEB-2002 (first entry)
 DT Listeria monocytogenes protein #316.
 DE Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
 KW vitamin B12; bacterial infection; disease.
 KW Listeria monocytogenes.
 OS W0200177335-A2.
 XX 18-OCT-2001.
 XX 11-APR-2001; 2001WO-FK01118.
 PF 11-APR-2000; 2000FR-0004629.
 PR (INSP) INST PASTEUR.
 XX Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;
 PI Dussurget O, Chetoui F, Nedjari H, Glaser P, Kunst F, Cossart P;
 PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
 PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
 PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
 PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
 PI Maduenio E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;
 PI Rose M, Voss H;
 XX WPI; 2002-010914/01.
 DR Genomic sequence for *Listeria monocytogenes*, useful e.g. for treatment
 PT and prevention of *Listeria* and related bacterial infections, and
 PT related polypeptides
 XX Claim 6; SEQ ID No 317; 192pp; French.
 PS The present invention relates to the genome sequence of *Listeria*
 CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
 CC it are useful for selecting probes and primers for detecting genes in *L.*
 CC monocytogenes and related organisms, and for studying genetic
 CC polymorphisms and other genomes. The present sequence is a protein

CC encoded by the genome sequence of the present invention. Proteins
 CC expressed from the genome sequence are useful for raising specific
 CC antibodies, identification of *L. monocytogenes* and related organisms, and
 CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
 CC B12. The genome sequence and proteins encoded by it are also useful for
 CC selecting compounds that regulate gene expression and cell replication
 CC and modulate *L. monocytogenes*-related diseases. In addition, the genome
 CC sequence and proteins encoded by it are useful in pharmaceutical and
 CC vaccine compositions for the treatment or prevention of infections by *L.*
 CC monocytogenes and related organisms.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1146 AA;
 Query Match 42.5%; Score 2494; DB 23; Length 1146;
 Best Local Similarity 45.7%; Pred. No. 1.le-176; Indels 30; Gaps 11;
 Matches 525; Conservative 186; Mismatches 408;
 QY 30 KKILVANGETAVRAFRALLETGAATVAIYPREDGSPHRSFASAEVRIEGSPVKAYL 89
 Db 5 KKVLANGETAIRVMRACTELKTKVAIYQEDTSGPHRYKSDEAYLVGAGKPIDAYL 64
 QY 90 DIDEIIGAACKVKDAIYPGYGFISENAQARECAENGITFIGPTPEVLDLTGDKSRV 149
 Db 65 DIENIEIAKESGDAIHGPGYGFISENIEFARCEQEGIIIEVGPCKSHLDMFGDKIR 124
 QY 150 AAKKAGLEPLAEST-PSKNIDDIYKSAEGQTYPIFVKAVAGGGGGRFVSPPDELK 208
 Db 125 QALLADIPVPGNSGPFVAGIKEVEEFGEKNGYPLMIKASLGGGGGRVVSKEHVESF 184
 QY 209 TEASREAEAAFGDGSVYVERAVINPOHIEVOILGDRTEVHLVERDCSLQRRQKV 268
 Db 185 ERASAEAAFGDNDEVYVEKCMNPKHIEVOILGDRTEVHLVERDCSLQRRQKV 244
 QY 269 APAQHLDELDRICADAVKFCRSIGYOGAGTVFLDEKGNHVFIEMNPRIOVEHTV 328
 Db 245 APCNAITSELNRICDAAVAKLMKNVDYINAGTVFLV-EGDDFYFIEVNPVQVEH 303
 QY 329 EVTEVDLVKAQMLAAGATLKEGLT---ODKIKTHGAALOCRTTTPNGFRPDGT 385
 Db 304 MITGIDIVQSQVLFADYALHDLQVAIPKQEDIHGSAIQSRITTEDPLNFMFPG 363
 QY 386 TAYRSPGAGVRLD-GAAQLGGEITAHFDSMLVMTKCGSDFETAVARAQALAEFTV 444
 Db 364 DTYRSTGGVRLDAGNGFGQTVVTTPFYDSLVLVLCVTMGMTFEQATKMRNLEFR 423
 QY 445 VATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPADDEQGRILDYLDVTVN 504
 Db 424 VKNTPFLLNVVRHDPDFASGNYNTSFDITTPLEKFKPHIRDGTKTLRYIGNVT 483
 QY 505 VRPKDVAAPIDKLPNIKDL-PP-PP-PP-PP-PP-PP-PP-PP-PP-PP-PP-PP-PP 560
 Db 484 IKHRD--KPVVAEPLPKIPYGSQISPTCKQILDAGPEGVVDVWKQEEVLLD 541
 QY 561 AHOSLLATRVRSFALKPAEAAVAKLTPELLSVEAGGATYDVAMRFLFEDPDRL 620
 Db 542 AHOSLLATRVRSKDFQVADAMAHLLPNMFSEFMGGATFDVAYFLNEDPVR 601
 QY 621 AMPNVNIOMLLRNTVCTYTPYDUSVCFKFAEASSGVDFIRFDALNDVSO 680
 Db 602 QIPNVFMOLLRGANAVGYNKPNVIREFVQSAQSGVDFVFDLSLWIKGMEVSD 661
 QY 681 VLEWTAVAEVAMAYSGDLSDPNEKLYDLYLKMABEIVKSGAHILAIDMAGLL 740
 Db 662 VREAG-KVVEAAICVYTGIDDDDTFTKYIDYKDMAKELVAGQTHILGIDMAG 720
 QY 741 VTKLVTALRRFDFLVVHHTHTAGGQALATYFAAAQAGADAVDAGSAPLSCT 800
 Db 721 AYRLGELKDTVDVPIHLHTHTSGNGITYTAAAVSAGVDIVDVASSAMSG 780
 QY 801 IVAAPFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTCGRVYR 860

Db 781 LYVGLVNGNRQTNLDNAQNSQIINHVEDVRHYKDFDNALNSPQTEVVIHEMPGQYTNL 840
Qy 861 RAQTALGLADRFELIENYAAVNMELGRPKVTPSSKVGDLALHLVAGVDPADPAD 920
Db 841 QQAIAVGLGRWDEVKEMTYVNMFGDIVKVTSSKVGDLALFMVONELSEEDVYEK 900
Qy 921 POKYDIPDSVIAFLRGELNPGGWPEPLRTRALEGRSEGRAPLTPVPEEQAHLDADDS 980
Db 901 GDTIDFDSVIEFFWGEIGQPVGPEKLOKLVLAGRT----PLTDRGALMEPVNFVDV 956
Qy 981 K-----ERNNSLNRLLFPKPTTEFLHRRRFGNTSALDREFFYGLVGEQRETLI 1029
Db 957 KAELEKMGYEPTKEDVISIYLPKVFLDYQDMINKYGDVTLDPTEYKGMRLGETIEV 1016
Qy 1030 RLPDVRTPLVRLDAISEDDKGMNVANVNGOIRPMVRDRSVESVTATAEKADSSNK 1089
Db 1017 ELEKGI-LLIKLNSIGEPIDGTRVIYFELNGQPREINQDMNVQSTVIARRKIDTTNP 1075
Qy 1090 GHVAAPFAG-VVTVTAEGDEVKAGDAVAIIEAMKMEATITASVDGKIERYVVVPAATKVE 1148
Db 1076 EHVGMTGSIQVIVVKKGDSVKKGDPLLLITEAMKMETIIQAPFDGEVSIYVSGDITIE 1135
Qy 1149 GDDLIVVVS 1157
Db 1136 SGDLLIEVN 1144

RESULT 13

AAU33972
ID AAU33972 standard; Protein; 1147 AA.

XX AAU33972;

AC AAU33972;

DT 14-FEB-2002 (first entry)

DE Staphylococcus aureus cellular proliferation protein #248.

XX Antisense; prokaryotic cellular proliferation protein;

XX antibiotic; antibacterial; drug design.

XX Staphylococcus aureus.

XX WO200170955-A2.

XX 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US09180.

XX 21-MAR-2000; 2000US-191078P.

XX 23-MAY-2000; 2000US-206848P.

XX 26-MAY-2000; 2000US-207727P.

XX 23-OCT-2000; 2000US-242578P.

XX 27-NOV-2000; 2000US-253625P.

XX 16-DEC-2000; 2000US-257931P.

XX 21-DEC-2001; 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

PI Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.

XX N-PSDB; AAS51831.

XX New polynucleotides for the identification and development of

PT antibiotics, comprise sequences of antisense nucleic acids -

XX Example 3; Seq ID No 5468; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to

CC prokaryotic cellular proliferation, their use in identifying the

CC genes, their use in the discovery of novel antibiotics, the essential

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

Db 705 PHILAKDMAGLLPKKAAAYELIGELKAAVDLPILHHTDTSNGLLITYKQADAGVDIID 764
 Qy 784 GASAPLSGTTQSPLSAIAVAFAHTRDRTGLSEAVSDLEPYWEAVRGLYLPFESGTPGP 843
 Db 765 TAVASMSGLTSQPSANSLYALNGFPRLHRTDIEGMSUSHYVSTVRYYSDFESDIKSP 824
 Qy 844 TGRVYRHEIFGGQSLNRAQATGALGLADRFELIEDNYAAVNEMLGRTKVTPTSSKVVGD 903
 Db 825 NTELYQHEMFGGQYNSLSQAQSLGLGERDEKVMYRRVNFGLGDIVKVTPTSSKVVGD 884
 Qy 904 ALHLVAGVADPAFAADPQRYDIPDSVIAFLRGELGNPFGWPEPLRTRALEGRSEKAP 963
 Db 885 ALYVQNDLDEQSVITDGYKLPFESVSPFKGEIGOPVNGFNKDLQAVLLKQGE---A 940
 Qy 964 LTVPEEQALDADDSKE-----RRNSLRLFPKPTPEFLHRRFRFGNTSAL 1012
 Db 941 LTRAPGEVLEPVDPEKVRLEEEEOQGPVTEQDIISVLYPKVYEQYIQTRNOYGNLSLL 1000
 Qy 1013 DDREFFYGLVGERETILRLPDVTRPLILRLDAISEDDKGMNVVANVANGQIRPMVRDR 1072
 Db 1001 DTPFTFFGMNGETVEIEI-DKGRLLIRLETISEPDENGNTIYYAMNGQARRIYKDE 1059
 Qy 1073 SVESVTATAEKADSSNGHVAAPFAGVVT-VTVAGDEGVKAGDAVAIEAMKMEATITAS 1131
 Db 1060 NVHTNANVKKPAKSNPSHIGAQMPSVTEKVSGETVKANQPLLITEAMKMETTIQAP 1119
 Qy 1132 VDGKIERVVVPAATKVEGGDLIVV 1156
 Db 1120 FDGVIKQVTNNGDITATGDLLEI 1144

RESULT 14

AAAG82677
 ID AAG82677 standard; Protein; 1151 AA.

AC AAG82677;

XX 03-SEP-2001 (first entry)

DE S. epidermidis open reading frame protein sequence SEQ ID NO:2448.

XX Staphylococcus epidermidis SRI strain; infection; diagnosis;

KW vaccination; endocarditis.

XX Staphylococcus epidermidis.

OS W0200134809-A2.

XX 17-MAY-2001.

XX 09-NOV-2000; 2000WO-US30782.

XX 09-NOV-1999; 99US-0164258.

PA (GLAX) GLAXO GROUP LTD.

XX KImmerly WJ;

XX WPI: 2001-316495/33.

DR N-PSDB; AAH53527.

XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,

PT useful for vaccinating against infections, e.g. endocarditis -

XX Claim 18; Page 644-645; 2188pp; English.

CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides

CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.

CC (I) and (II) can have antibacterial activity and therefore can be used

CC in vaccination. The nucleic acids (I) may be used to produce the

CC S. epidermidis polypeptides (II) via the production of vectors

CC containing them which are used to produce hosts cells which express the

CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be

CC used to vaccinate subjects and to raise antibodies against the bacteria.
 CC The polypeptides may also be used to assay for other inhibitors of their
 CC activity and therefore identify compounds that may be used for the
 CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
 CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
 CC polynucleotide sequences from the present invention. AAH55091 to
 CC AAH55098 represent oligonucleotide sequences and primers which are used
 CC in the exemplification of the present invention.
 CC N.B. The present invention specifically claims all the polynucleotide
 CC sequences given in the sequence listing of the present specification,
 CC however the sequence listing only goes up to SEQ ID NO:4454 so even
 CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
 CC no sequences are present for SEQ ID NO:4455 to 4464.

XX SQ Sequence 1151 AA;

Query Match 42.0%; Score 2465; DB 22; Length 1151;

Best Local Similarity 45.4%; Pred. No. 1.6e-174;

Matches 524; Conservative 198; Mismatches 399; Indels 34; Gaps 14;

Qy 26 LPAFKKILVANGETAVRAFRALLETGAATVAIYPRDRGSHRSFASEAVRIGTESPV 85

Db 5 LKQIKLLLVANGETAIRIFRAAAELNISTVAIYSNEDKSLHRYKADESIVLSGDLGPA 64

Qy 86 KAYLDEIIEGAARKVKADAIYPGYFLSENAQLARECAENGITFTGPTPEVLDTGDKS 145

Db 65 ESYLNIERIIEVALRAGVDAIHPGYGFLSENEQFAARRCAEEGKIFGPHLEHLMFQDKV 124

Qy 146 RAVTAAKRAKGLPVL-AESTPSKNIDIVKSAEGQTYPIFVKAVAGGGMRFRVSSDEL 204

Db 125 KARTTAIINANLPVPGTDGPISFEAAEQFANEAGYPLMIKATSGGGKGMRIRESSEL 184

Qy 205 RKLATEASKEAAAFPGDGVSVYVERAVINPOHIEVQLGDRTEVVVHLYERDCSLQRHQK 264

Db 185 EDAPHRASEAKESGSEVVIERYIDNPKHIEVQVIGVEFGNIHLHYERDCSVQRHQK 244

Qy 265 VVEIAPAOHLDELDRICADAVKFCISGYOGATVEFLV--DEKGNHVIENNPRTQV 322

Db 245 VVEVAPSVGLSKLREICDAAIQLMENIKYVNAVTEFLVSGDE---FFIEVNPVQV 301

Qy 323 EHTVTEEVTEVDLVKAQMLAAGATL---KELGLTQDKIKTHGAALOCRIITPEDNNGFR 379

Db 302 EHTITMITGIDIVKTQILVADGESLFGDKISMPQONEIQTLGYAIOCRITTEPTDNFM 361

Qy 380 PDTGTITAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVMKTCRGSDFETAVARAQALA 438

Db 362 PDSGTIIAYRSRGGVGVRLDAGDGFQGAIEISPYDLSLLVKLSTHAVSFQAEKMERSLR 421

Qy 439 EFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLDV 498

Db 422 EMRIRGVKTNIPFLINVRNDRKFRSGDYTTKFIETPELFDIAPTLDRGTKTLEYIGNVT 481

Qy 499 VNKPHGV--RPKDVAAPIDKLPNIKDLPLPR--GSRDLKQLGPAFAFARDLREQDALAVT 554

Db 482 INGFNVKRPKP-EYESTKIPKISOKKINQLFGTKLEQHGPTGVTNVRREQEDVLIT 540

Qy 555 DTTFRDAHQSLLATRVRSFALKPAEAAVAKLTPELLSVSEAMGATYDVMRFLFEDPWR 614

Db 541 DTTFRDAHQSLLATRVRTKDMNIAASKTAEVFKDSFSLMMGGATFDVAVNFKENPWR 600

Qy 615 LDELREAMPNVIOMLLRGRTNVGYTPYDSCVAFYKAAASSGVDFRIFEDALNDVDSOM 674

Db 601 LERLKAIPNLVLFQMLLRASNAVGYKPNVIVKFVHESAKAGVDVFRIFDSLNWDDM 660

Qy 675 RPAIDAVLETNATAEAVAMAYSGD-LSDPNEKLYTLDYLLKMAEETVKSAGHIAIKDMA 733

Db 661 KVANAQVEAG-MVSEGTICYTGDIILNAERSNIYTDYVVKMALEREGPHILAIKDMA 719

Qy 734 GLLRPAATKLVTLRRREFDLPVHVHTHTDAGGOLATYFAAAQAGADAVDGASAPLSGTT 793

Db 720 GLLKPKAAEYELIGELREATHLPILHHTDTSNGLLTYKQADAGVDIIDTAVASMSGLT 779

Qy 794 SQPSLSAIVAAFAHTRDRTGLSEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIP 853

Db 780 SOPSANSLYALNGPPRLNRITDIDGLELSHYWSVVRPYIADFDSDIKSPNTEIYQHEMP 839
QY 854 GGQSLNLRQAATAGLADRFELIEDNYAAVNEMLGRPKVTSSKVVGDALHLVYAGVD 913
Db 840 GGQYSLNLSQAQSLGLGERFDEVKEMRYRNFLGDLVKVTSSKVVGDALVYQNDLD 899
QY 914 PADFAADPOKYDIPDSVIAFLRGELNPPGWPPEPLRTRALGRSEGRAPLTPVPEEQA 973
Db 900 EDTVINDGVKLDPEVSVYFFKGDIGQPVNGFNKKLQDVLKGOQ----PITERGLEYE 955
QY 974 HUDAD-----DKERNNSLRLFFPKPTTEEFLEHRRFRGNTSALDREFFYGLV 1022
Db 956 PVDFAIRQELSDIOQDEVTQDIISYVLYPKVYQYIQTKQFQGVNSLLDPTFLFGMR 1015
QY 1023 EGRETLRLPDRVTLLVRLDLAISPDGKGMNVANVNGQIRPMVRDRSVESVTAAE 1082
Db 1016 NGETVEIEI-DTGKRLIILKLETISEPDENKRTIYAMNQARRIYIODENKVTNANVKP 1074
QY 1083 KADSSNKGHVAAFPAGVVT-VTVAGDEVKAGDAVAIIEAMKMEATITASVDGKIERYVV 1141
Db 1075 KADKSNPHIGAQMPSGVTEVKVSVGDEVQANQPLLITEAMKMETIQAPDGLIKQINV 1134
QY 1142 PAATKVEGGDLIVV 1156
Db 1135 ANGDAIATGDLLEVEI 1149

RESULT 15

ABP38583
ID ABP38583 standard; Protein; 1154 AA.
XX
AC ABP38583;
XX
DT 24-JUL-2002 (first entry)
XX
DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3428.
XX
KW Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
KW antibacterial; gene therapy.
XX
OS Staphylococcus epidermidis.
XX
PN USG380370-B1.
XX
PD 30-APR-2002.
XX
PF 13-AUG-1998; 98US-0134001.
XX
PR 14-AUG-1997; 97US-055779P.
XX
PR 08-NOV-1997; 97US-064964P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Doucette-Stamm LA, Bush D;
XX
XX WPI; 2002-381255/41.
DR N-PSDB; ABN91128.
XX
PT Novel isolated nucleic acid encoding a Staphylococcus epidermidis
PT polypeptide, useful for diagnosing and treating bacterial infections -
XX
XX Disclosure; SEQ ID 3428; 267pp; English.
XX
CC ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC given in ABP35124 to ABP37960. The S. epidermidis sequences have
CC antibacterial activity and can be used in gene therapy. The sequences
CC can also be used in the diagnosis and treatment of bacterial infections,
CC particularly S. epidermidis infections. The sequences can be used to
CC screen for compounds able to interfere with the S. epidermidis life
CC cycle or inhibit S. epidermidis infection.
CC N.B. The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from the
CC USPTO web site.
XX
SQ Sequence 1154 AA;
Query Match 42.0%; Score 2465; DB 23; Length 1154;
Best Local Similarity 45.4%; Pred. No. 1.6e-174;
Matches 524; Conservative 198; Mismatches 399; Indels 34; Gaps 14;
QY 26 LPAPKILVANGETAVRAAFRAAETGAATVAIYIPREDRSGFHRFASAVRIGTEGSPV 85
Db 6 LKQIKLLVANGETAIRIFRAAELINISTVAIYNSNEDKSLHRYKADESILVSGDLGPA 65
QY 86 KAYLOIDEIIGAAKVKADAIYGVGFLSENQAOLARECAENGITFTGTPPEVLDLTGDKS 145
Db 66 ESYLNERIEIIVALRAGVDIAHPGYGFLSENEQFARCAEIGIKFIGPHLEHLDMPGDKV 125
QY 146 RAVTAAKKAGLPLV-AESTPSKNIDDIKSAEQTYPIFVKAVAGGGGRMFVSSPDEL 204
Db 126 KARTTAINANLPVPGTDGPIESFEAAEQFANEAGVPLMIKATSGGGGKMRIVRESSEL 185
QY 205 RKLATEASREAAFGDSVYVERAVINPOHIEVQLIGDRTGEVHLYERDCSLORRHK 264
Db 186 EDAPFRKSEAEKSEFNGSEVYIERYIDNPKHIEVQVIGDEFGNIIHLYERDCSVORRHK 245
QY 265 VVEIAPAQHLDPEDLRICADAVKFCRSIGYOGAGTVFVLV--DEKGNHVFTEMPIRQV 322
Db 246 VVEVAPSVGLSNKLRERICDAAIQIMENIKYVNAVGTVEFLVSGDE---FFFIEVNPVQV 302
QY 323 EHTVTEVEVDLVKAQMLAAAGATL---KELGLTQDKIKTHGAALQCRITTEDPNNGPR 379
Db 303 EHTITEMITGIDIVKTQILVANGESLFGDKISMPQONEIQTILGYAQCRITTEDPTNDFM 362
QY 380 PDTGTITAYRSGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQALA 438
Db 363 PDGSGTIIAYRSSGGFGVRLDAGDGFQGAIEISPYDSLLVKLTHAVSFQAEKEMERSUR 422
QY 439 EFTVSGVATNIGLRALLREEDFTSKRIATFGIDHPHLLQAPPADDEQGRILDYLDVY 498
Db 423 EMRIGVKTNIPFLINVRNDKFRSGDYTKFIEETPELFDIAPTLDRGKTLEYIGNVT 482
QY 499 VNKPHGV--RPKDVAAPIDKLPNIKDLPLR--GSRDLKQLGPAFAFARDLREQDALAVT 554
Db 483 INGFNPVEKRPKP-EVESTKIPKISQKINQLFGTKQILEQHGPTGVTNWVREQEDVLIT 541
QY 555 DTFERDAHQSLLATVRSFALKPAAEAVAKLTPELLSVEAWGATYDVAMRFLFEDPWDR 614
Db 542 DTFERDAHQSLLATVRRTKDMNIAASKTAEVFKDSFSELMWGGATFDVAYNFKENPWR 601
QY 615 LDELREAMPNVNIOMLLRGRTVGYTPYDPSVCRAFVKEAASGVDIFRIFDALNDVDSOM 674
Db 602 LERLKAIPNVLFQMLLRASNAVGYKNYPDNVIKFFVESAKAGVDVFRIFSLNWDQM 661
QY 675 RPAIDAVLETNTAVAEAMAYSGD-LSDPNEKLYTLDYLYLKMAEEIVKSGAHILAIKDMA 733
Db 662 KVANEAVQEG-MVSEGTICYTGDILNACRSNIYTLDYLYVVKMAKELEREGFHILAIKDMA 720
QY 734 GLLRPAAVTKLVTLALRREFDLPVHVHTHTAGCOLATYFAAAQAGADAVDGSAPLSGTT 793
Db 721 GLLKPKAAVYELIGELREATHPLPHLHTHTDTSGLLTYYKQAIADAGVDIIDTAVASMSGLT 780
QY 794 SOPSLSAAVAFAHTRRDTGLSEAVSDLEVPWEAVRGYLPFESGTPGPTGRVYRHEIP 853
Db 781 SOPSANSLYALNGPPRLNRITDIDGLELSHYWSVVRPYIADFDSDIKSPNTEIYQHEMP 840
QY 854 GGQSLNLRQAATAGLADRFELIEDNYAAVNEMLGRPKVTSSKVVGDALHLVYAGVD 913
Db 841 GGQYSLNLSQAQSLGLGERFDEVKEMRYRNFLGDLVKVTSSKVVGDALVYQNDLD 900
QY 914 PADFAADPOKYDIPDSVIAFLRGELNPPGWPPEPLRTRALGRSEGRAPLTPVPEEQA 973
Db 901 EDTVINDGVKLDPEVSVYFFKGDIGQPVNGFNKKLQDVLKGOQ----PITERGLEYE 956

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2003, 18:59:26 ; Search time 17 seconds
(without alignments)
2002.489 Million cell updates/sec

Title: US-09-974-973-2
Perfect score: 5865
Sequence: 1 MTALTIGLLKGIITLVST.....RVVVPAAKVEGGDLIVVVS 1157

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5759	98.2	1140	4	US-09-220-081-2
2	5759	98.2	1140	4	US-09-677-575-2
3	2465	42.0	1154	4	US-09-134-001C-3428
4	1192	20.3	694	4	US-09-433-043B-126
5	1039.5	17.7	593	4	US-09-433-043B-122
6	1032.5	17.6	447	1	US-08-611-107-6
7	1032.5	17.6	447	2	US-08-422-560A-6
8	1032.5	17.6	447	4	US-08-468-793-6
9	1005.5	17.1	453	2	US-08-611-107-8
10	1005.5	17.1	453	2	US-08-422-560A-8
11	1005.5	17.1	453	4	US-08-468-793-8
12	1003.5	17.1	453	4	US-09-433-043B-121
13	996.5	17.0	453	1	US-07-956-700B-6
14	996.5	17.0	453	1	US-08-476-537-6
15	996.5	17.0	453	1	US-08-485-607-6
16	996.5	17.0	453	2	US-08-475-879-6
17	996.5	17.0	453	4	US-09-433-043B-6
18	954	16.3	605	4	US-09-433-043B-123
19	940	16.0	448	1	US-08-074-121-3
20	940	16.0	448	5	PCT-US94-06447-3
21	933	15.9	427	1	US-07-956-700B-3
22	933	15.9	427	1	US-08-476-537-3
23	933	15.9	427	1	US-08-485-607-3
24	933	15.9	427	2	US-08-475-879-3
25	933	15.9	427	4	US-09-433-043B-3
26	928	15.8	536	4	US-08-662-344-2
27	921	15.7	449	1	US-08-074-121-6

28	921	15.7	449	5	PCT-US94-06447-6
29	849	14.5	461	4	US-09-134-001C-3604
30	842.5	14.4	722	4	US-09-433-043B-125
31	786	13.4	703	4	US-09-433-043B-124
32	782	13.3	411	4	US-09-134-001C-3221
33	613	10.5	853	4	US-09-433-043B-119
34	605	10.3	852	4	US-09-433-043B-118
35	593.5	9.9	2257	1	US-08-611-107-10
36	583.5	9.9	2257	2	US-08-422-560A-10
37	583.5	9.9	2257	4	US-08-468-793-10
38	561.5	9.6	2325	3	US-08-417-089-6
39	561.5	9.6	2325	4	US-08-695-651-6
40	561.5	9.6	2325	4	US-08-930-285-6
41	561.5	9.6	2325	4	US-08-695-421-6
42	561.5	9.6	2325	4	US-08-697-826A-10
43	559	9.5	2172	1	US-08-611-107-31
44	555.5	9.5	2254	2	US-08-677-010-3
45	555.5	9.5	2254	2	US-08-790-519-3

ALIGNMENTS

RESULT 1

US-09-220-081-2
; Sequence 2, Application US/09220081
; Patent No. 6171833
; GENERAL INFORMATION:
; APPLICANT: Sinskey, Anthony J.
; APPLICANT: Lessard, Philip A.
; APPLICANT: Willis, Laura B.
; APPLICANT: Stephanopoulos, Gregory
; TITLE OF INVENTION: Pyruvate Carboxylase from Corynebacterium glutamicum
; FILE REFERENCE: 1533.0790000
; CURRENT APPLICATION NUMBER: US/09/220,081
; CURRENT FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1140
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-220-081-2

Query Match	98.2%	Score	5759	DB	4	Length	1140
Best Local Similarity	99.4%	Pred. No.	0				
Matches	1133	Conservative	5	Mismatches	2	Indels	0
Gaps	0						
Qy	18	VSTHTSSTLPAPFKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSEAVR	77				
Db	1	MSTHTSSTLPAPFKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSEAVR	60				
Qy	78	IGTEGSPVKAYLDIDEITGAARKVKADAIYPGYFLSNAQLARECAENGITFTGTPTEV	137				
Db	61	IGTEGSPVKAYLDIDEITGAARKVKADAIYPGYFLSNAQLARECAENGITFTGTPTEV	120				
Qy	138	LDLTGDKSRVTAATAKAGLPVLAESTPSKNIDDIIVKSEGTYPYFVKAVAGGGGRGMR	197				
Db	121	LDLTGDKSRVTAATAKAGLPVLAESTPSKNIDDIIVKSEGTYPYFVKAVAGGGGRGMR	180				
Qy	198	VSSPDLEKRLATEASREAAAFDGGSVYVERAVINPQIHIEVQILGDRTEGVVHLYERDCS	257				
Db	181	VASPDLEKRLATEASREAAAFDGGSVYVERAVINPQIHIEVQILGDRTEGVVHLYERDCS	240				
Qy	258	LQRHQKVVETAPQHLDPDLDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEMN	317				
Db	241	LQRHQKVVETAPQHLDPDLDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEMN	300				
Qy	318	PRIOQVHTVEEVTVDLVKQMLAAGATLKLGLTODKIKTHCAALOCRTITTEDPNN	377				
Db	301	PRIOQVHTVEEVTVDLVKQMLAAGATLKLGLTODKIKTHCAALOCRTITTEDPNN	360				
Qy	378	FRPDTGTTITVRSPPGAGVRLDGAALQGLGEITAHFDSMLVKMTCRGSDFETAVARAORAL	437				

```
Db 361 FRPDTGCTITAYRSPGGAGVRLDGAALGGITAHFDSMLVKWTCRGSDFETAVARAQAL 420
QY 438 AEFVTSVATNIGFLRALLREEDFTSKRIATGFIQDHPHLLQAPPADDEQGRILLYLADV 497
Db 421 AEFVTSVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILLYLADV 480
QY 498 TVNKPVGVRPKDVAAPIDKLPNTKIDLPGRGSRDLKQLGPAAFARDLRQDALAVTDIT 557
Db 481 TVNKPVGVRPKDVAAPIDKLPNTKIDLPGRGSRDLKQLGPAAFARDLRQDALAVTDIT 540
QY 558 FRDAHQSLLATRVRSFALKPAAEAVAKLTPPELLSVKAWGATYDVAMRFLFEDPDWRDLDE 617
Db 541 FRDAHQSLLATRVRSFALKPAAEAVAKLTPPELLSVKAWGATYDVAMRFLFEDPDWRDLDE 600
QY 618 LREAMPNVNTOMLLRGRNTVGYTPYSDVCRAPVKEAASGGVDIFRIFDALNDVSQMRPA 677
Db 601 LREAMPNVNTOMLLRGRNTVGYTPYSDVCRAPVKEAASGGVDIFRIFDALNDVSQMRPA 660
QY 678 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYLLKMAEEIVKSGAHILAIDKMGALLR 737
Db 661 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYLLKMAEEIVKSGAHILAIDKMGALLR 720
QY 738 PAAVTKLVTLARREFDLPVHVHTDHTAGGOLATYFAAAQAGADAVDGSAPLSGTTSQPS 797
Db 721 PAAVTKLVTLARREFDLPVHVHTDHTAGGOLATYFAAAQAGADAVDGSAPLSGTTSQPS 780
QY 798 LSAIVAAFAHTRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGOL 857
Db 781 LSAIVAAFAHTRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGOL 840
QY 858 SNLRAGATLGLADREFELTDYNAAVNMLGRPTKVTPESSKVVGDALHLVAGVDPADF 917
Db 841 SNLRAGATLGLADREFELTDYNAAVNMLGRPTKVTPESSKVVGDALHLVAGVDPADF 900
QY 918 ADPOKYDIPDSVIAFLRGELNPPGWPPELTRALEGRSEKAPLTVPEEEQAHLDA 977
Db 901 ADPOKYDIPDSVIAFLRGELNPPGWPPELTRALEGRSEKAPLTVPEEEQAHLDA 960
QY 978 DSKERNNSLNRLLFPKPTPEEFLEHRRRGNTSALDDREFFYGLVGEGRETLRLPDVTRP 1037
Db 961 DSKERNNSLNRLLFPKPTPEEFLEHRRRGNTSALDDREFFYGLVGEGRETLRLPDVTRP 1020
QY 1038 LLVRLDAISEPDDKGRNVVANYNGQIRPMVRDRDSVESVTATAEKADSSNKGHVAAPFA 1097
Db 1021 LLVRLDAISEPDDKGRNVVANYNGQIRPMVRDRDSVESVTATAEKADSSNKGHVAAPFA 1080
QY 1098 GVVTVTVAEGDEVKAGDAVAIIEMKMEATITASVDGKIERYVVPVPAATKVEGGDLIVVVS 1157
Db 1081 GVVTVTVAEGDEVKAGDAVAIIEMKMEATITASVDGKIERYVVPVPAATKVEGGDLIVVVS 1140
```

RESULT 2

```
US-09-677-575-2
: Sequence 2, Application US/09677575
: Patent No. 6403351
: GENERAL INFORMATION:
: APPLICANT: Sinskey, Anthony J.
: APPLICANT: Lessard, Philip A.
: APPLICANT: Willis, Laura B.
: APPLICANT: Stephanopoulos, Gregory
: TITLE OF INVENTION: Pyruvate Carboxylase from Corynebacterium glutamicum
: FILE REFERENCE: 1533.0790000
: CURRENT APPLICATION NUMBER: US/09/677,575
: PRIOR FILING DATE: 2000-10-03
: PRIOR FILING DATE: 1998-12-23
: NUMBER OF SEQ ID NOS: 36
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 2
: LENGTH: 1140
: TYPE: prt
: ORGANISM: Corynebacterium glutamicum
```

```
US-09-677-575-2
Query Match 98.2%; Score 5759; DB 4; Length 1140;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1133; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
QY 18 VSTHTSSLPAPKTLVANRGEIAVRAFAALETGAATVATVYIPREDRSGSPHRSFASAVR 77
Db 1 MSTHTSSLPAPKTLVANRGEIAVRAFAALETGAATVATVYIPREDRSGSPHRSFASAVR 60
QY 78 ICTESPVKAYLDIDEIIGAARKVADAIYCYGFLSENAQLARECAENGITFTGPTPEV 137
Db 61 ICTESPVKAYLDIDEIIGAARKVADAIYCYGFLSENAQLARECAENGITFTGPTPEV 120
QY 138 LDLTGDKSRVATAAKKAGLPVLAESTPSKNIDDIKVSAGQTYPIFKVAVAGGGGRMRF 197
Db 121 LDLTGDKSRVATAAKKAGLPVLAESTPSKNIDDIKVSAGQTYPIFKVAVAGGGGRMRF 180
QY 198 VSSDELKRLKLAESRAEAAFGDGVVVERAVINPOHIEVQIILGDRGTGEVHVLYERDCS 257
Db 181 VASDELKRLKLAESRAEAAFGDGVVVERAVINPOHIEVQIILGDRGTGEVHVLYERDCS 240
QY 258 LQRHQKVVETAEVLDLVKAQMLRAGATLKLGLTQDKIKTHGAALOCRTTDEPNNG 317
Db 241 LQRHQKVVETAEVLDLVKAQMLRAGATLKLGLTQDKIKTHGAALOCRTTDEPNNG 300
QY 318 PRIOVEHTVTEVEVDLVKAQMLRAGATLKLGLTQDKIKTHGAALOCRTTDEPNNG 377
Db 301 PRIOVEHTVTEVEVDLVKAQMLRAGATLKLGLTQDKIKTHGAALOCRTTDEPNNG 360
QY 378 PRPDGTITAYRSPGGAGVRLDGAALGGITAHFDSMLVKWTCRGSDFETAVARAQAL 437
Db 361 PRPDGTITAYRSPGGAGVRLDGAALGGITAHFDSMLVKWTCRGSDFETAVARAQAL 420
QY 438 AEFVTSVATNIGFLRALLREEDFTSKRIATGFIQDHPHLLQAPPADDEQGRILLYLADV 497
Db 421 AEFVTSVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILLYLADV 480
QY 498 TVNKPVGVRPKDVAAPIDKLPNTKIDLPGRGSRDLKQLGPAAFARDLRQDALAVTDIT 557
Db 481 TVNKPVGVRPKDVAAPIDKLPNTKIDLPGRGSRDLKQLGPAAFARDLRQDALAVTDIT 540
QY 558 FRDAHQSLLATRVRSFALKPAAEAVAKLTPPELLSVKAWGATYDVAMRFLFEDPDWRDLDE 617
Db 541 FRDAHQSLLATRVRSFALKPAAEAVAKLTPPELLSVKAWGATYDVAMRFLFEDPDWRDLDE 600
QY 618 LREAMPNVNTOMLLRGRNTVGYTPYSDVCRAPVKEAASGGVDIFRIFDALNDVSQMRPA 677
Db 601 LREAMPNVNTOMLLRGRNTVGYTPYSDVCRAPVKEAASGGVDIFRIFDALNDVSQMRPA 660
QY 678 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYLLKMAEEIVKSGAHILAIDKMGALLR 737
Db 661 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYLLKMAEEIVKSGAHILAIDKMGALLR 720
QY 738 PAAVTKLVTLARREFDLPVHVHTDHTAGGOLATYFAAAQAGADAVDGSAPLSGTTSQPS 797
Db 721 PAAVTKLVTLARREFDLPVHVHTDHTAGGOLATYFAAAQAGADAVDGSAPLSGTTSQPS 780
QY 798 LSAIVAAFAHTRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGOL 857
Db 781 LSAIVAAFAHTRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGOL 840
QY 858 SNLRAGATLGLADREFELTDYNAAVNMLGRPTKVTPESSKVVGDALHLVAGVDPADF 917
Db 841 SNLRAGATLGLADREFELTDYNAAVNMLGRPTKVTPESSKVVGDALHLVAGVDPADF 900
QY 918 AADPOKYDIPDSVIAFLRGELNPPGWPPELTRALEGRSEKAPLTVPEEEQAHLDA 977
Db 901 AADPOKYDIPDSVIAFLRGELNPPGWPPELTRALEGRSEKAPLTVPEEEQAHLDA 960
QY 978 DSKERNNSLNRLLFPKPTPEEFLEHRRRGNTSALDDREFFYGLVGEGRETLRLPDVTRP 1037
Db 961 DSKERNNSLNRLLFPKPTPEEFLEHRRRGNTSALDDREFFYGLVGEGRETLRLPDVTRP 1020
```

Qy 555 DTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDR 61

```

Qy 31 KILVANGETIAVRAALETGAATVAIYPREDSGPHRSFASAEVRICTEG--SPVKAY 88
Db 21 KILVANGETIPIRIFRTAHELMSQTAIYSHEDLSLTHKQKADAEVYIGVGQYTPVGAY 80
Qy 89 LDIDEITGAKKVADAIYPGYGLSENQAOLARECAENGITFIGPTPEVLDLTGDKSRV 148
Db 81 LAIDEITGAQKHVDITPHGYGLSENSEFADKVKVAGITWIGGPAEVIDSGDKVSAR 140
Qy 149 TAAKAGLPVLAEST-PSKNIDDIVKSAEGOTYPIFKAVAGGGGGRMRFVSSPDELRL 207
Db 141 NLAAKANVTYVPGTPEIEVEEALDFVNEYGYPIVKAAGGGGGRMRFVSSPDELRL 200
Qy 208 ATEASREAEAFGSGSVYVERAVINPOHIEVQILGDRTEGVVHLYERDCSLQRHOKVVE 267
Db 201 FORATSEARTAGFNGTCFVERFDKPKHIEVQLLADNHNHNVHLFERDCSVQRHOKVVE 260
Qy 268 IAPAQHLDPELDRICADAVKFCRSIGYOGAGTVFEVLDEKGNHVFTEMNPRIOQVHTVT 327
Db 261 VAPAKTLPREVDRDAILTDVAKLAKCEGYRNAGTAELFLVDNQRHVFIEINPRIOQVHTVT 320
Qy 328 EEVTEDVLVKAOMLAAGATLKEGLTQDKIKTHGAALOCRIITTEDPNNNGFRPDTGITA 387
Db 321 EETIGDIDVAAQIQ-AAGASLPQLGLFQDKITTRGFALQCRITTEDPAKNFQPDGTGRIEV 379
Qy 388 YRSPGGAGVRLDGA-AOLGEITAHFDSMLVKMTCRGSDFETAVARAORALAEFTVSGVA 446
Db 380 YRSAGNGVRLDGNVAGYTIISPHYDSMLVKCSGSTEIVRRKMIRALIEFRINGVK 439
Qy 447 TNIGFLRALLREEDFTSKRTATGFIGDHPHLLQAPPADDEQGRILDYLDADVTYVKNKHGVR 506
Db 440 TNIFELTLTNVFIETGVTGTFIDTQPFQVNSONRAQLLHLADVA-----491
Qy 507 PKOVAAPIDKLPNIKDLPLRGSRLKQLGPAFAFARDLREQDALAVDTTFRDAHOSLL 566
Db 492 -----491
Qy 567 ATRVRSFALPAEAVAKLTPELLSVEAWGATYDVAMREFEDPMDRLDELREAMPNVN 626
Db 492 -----491
Qy 627 IQMLLRGNTVGTYPDSVCRAPFVKEAASSGVDFIRFDALNDVDSQMRPAIDAVLETNT 686
Db 492 -----491
Qy 687 AVAEVAMAYSGDLSDPNEKLYTLDYLLKMAEIVKSGAHLIAIKDMAGLLRPAAVTKLVT 746
Db 492 -----491
Qy 747 ALREFDLPVHVHTHTAGQLATYFAAQAGADAVDASAPLSGTTSPQSLSAIYAFA 806
Db 492 -----491
Qy 807 HTRDRTGLSEAVSDLEPYEAVNGLYLPFESGTPGTGRVYRHEITPGGOLSLMRAQATA 866
Db 492 -----491
Qy 867 LGLADREFLEDNVAANEMLGRP-TKVTPSSKWVGDLALHLVGAGVDPAADAPQKYD 925
Db 492 -----DNGSSIKGQIGLPLKSNPSV-----EHSYN 517
Qy 926 IPDSVIAFLRGELNPGGWPPEPLRTRALEGRSEKAPLTVPEPEEQAHLDADDKERRN 985
Db 518 -----517
Qy 986 SLNRLLPKPTPEFLEHRRRRFGNTSALDDREFFYGLVEGRETILRLPDVTRPLLVRLDAI 1045
Db 518 -----MYPRVVEDFQKMYRETYGDLVLPTRSFSLPSETDEIEVWIEQGT-LIILQAV 571
Qy 1046 SEPDDK-GMNVYVANVGOIRPMRVDRSVESVTATAEKADSSNKGHVAAPFAGV-VTVT 1103
Db 572 GDLNKKTGEREVFDLNGEMRKIRVADRSQKVETVTKSKADMHDPLHIGAPMAGVIVEVK 631

```

RESULT 6

```

Qy 1104 VAEGDEVKAGDAVALIEAMKMEATITASVDGKIERYVVPVPAATKVEGGDLIVVV 1156
Db 632 VHKGSLIKKGQPVAVLSAMKMEMIISPSDQGVKEVFSVDGENVDSSOLLVILL 684

RESULT 5
US-09-433-043B-122
; Sequence 122, Application US/09433043B
; Patent No. 6399342
; GENERAL INFORMATION:
; APPLICANT: HASELKORN, ROBERT
; TITLE OF INVENTION: CYANOBACTERIAL AND PLANT ACETYL-CoA CARBOXYLASE
; FILE REFERENCE: ARCD:338US
; CURRENT APPLICATION NUMBER: US/09/433,043B
; CURRENT FILING DATE: 1999-10-25
; PRIOR APPLICATION NUMBER: 08/475,879
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 07/956,700
; PRIOR FILING DATE: 1992-10-02
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 122
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-09-433-043B-122

Query Match 17.7%; Score 1039.5; DB 4; Length 593;
Best Local Similarity 44.3%; Pred. No. 6.3e-74;
Matches 229; Conservative 80; Mismatches 177; Indels 31; Gaps 6

Qy 29 FKILVANGETIAVRAALETGAATVAIYPREDSGPHRSFASAEVRICTEGSPVKAY 88
Db 3 FDKILIANGETIALRLRACEEMGIATIAVSTVDNRNALHVQLADEAVCIG-EPASAKSY 61
Qy 89 LDIDEITGAKKVADAIYPGYGLSENQAOLARECAENGITFIGPTPEVLDLTGDKSRV 148
Db 62 LNIPIIAAALTRNASAIHPGYGLSENKAEICADHIAFIGTPEAIRLMGDKSTAK 121
Qy 149 TAAKAGLPVLAES-----TPSKNIDDIVKSAEGOTYPIFKAVAGGGGGRMRFVSSPDE 203
Db 122 ETMOKAGVPTVPGSEGLVETEQEGLE---LAKDIGYPMIKATAGGGGRMRLVRSPE 177
Qy 204 LRKLATEASREAEAFGSGSVYVERAVINPOHIEVQILGDRTEGVVHLYERDCSLORHO 263
Db 178 FVKFLAQGEAGAFGNAGVTIEKFIERPHIEFQILADNVGNVHILGERDCSLQRNQ 237
Qy 264 KVVETAPQHLDPDLDRICADAVKFCRSIGYOGAGTVFEVLDEKGNHVFTEMNPRIOVE 323
Db 238 KULEEAPSPALDSDLERKMGMQAAVRAAOFINYAGAGTIEFLDRSGGQYFMEMNTRIOVE 297
Qy 324 HVTVEVTEVDLVKAOMRLAAGATLKEGLTQDKIKTHGAALQCRITTEDPNNNGFRPDTG 383
Db 298 HVTTEMTVGVLDLLEQIRIAQGERLR--LTQDQVYLRGHAIECRINAEDPDHOFPRPAG 354
Qy 384 TITAYRSPGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDFETAVARAORALAEFTVS 443
Db 355 RYSGVLPFGGCVGRIDSHVYTDYQIPPYDSLSLIGKLVWGPDRATAINRMKRALRECAIT 414
Qy 444 GVATNIGFLRALLREDEFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLA-----495
Db 415 GLPTTIGFHQRIMENPQFLOGNVSTSFVOE---MNKPLDFNEIROLLTTIAQTDAIEVT 470
Qy 496 -----DVTNKPCHVRPKDVAAPIDKLPNKKDLPLP 526
Db 471 LKSDDFELTVKRVKGVNNSVVPVVTAPLSGVVSGCLP 507

RESULT 6

```

```

US-08-611-107-6
; Sequence 6, Application US/08611107
; Patent No. 5801233
; GENERAL INFORMATION:
; APPLICANT: Haselkorn, Robert
; APPLICANT: Gornicki, Piotr
; TITLE OF INVENTION: NUCLEIC ACID COMPOSITIONS ENCODING
; TITLE OF INVENTION: ACETYL-CoA CARBOXYLASE AND USES
; TITLE OF INVENTION: THEREFOR
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/611,107
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US SN 07/956,700
; FILING DATE: 02-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US93/09340
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US SN 08/422,560
; FILING DATE: 14-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: ARCD:221
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 447 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
;
US-08-611-107-6
Query Match 17.6%; Score 1032.5; DB 1; Length 447;
Best Local Similarity 48.2%; Pred. No. 1.4e-73;
Matches 216; Conservative 71; Mismatches 148; Indels 13; Gaps 4;

QY 29 FKKILVANGETIAVRAALETGAATVAIYPREDGSGFHRSEAVRIGTGSPVKAY 88
Db 3 FDKILLANGETIALRLTRACEEMGIATIAVSHVTDNRNALHVQLADEAVCGIG-EPASAKSY 61

QY 89 LDIDETIGAAGKVKADAIYPGYGFLSENALARECAENGITFTGPTPEVLDLTGDKSRV 148
Db 62 LNIPIIAAALTRNASHAIHPGYGFLSENALARECAENGITFTGPTPEVLDLTGDKSRV 121

QY 149 TAAKAGLPLVAES-----TPSKNIDDIVKSAEGQTYPIFVKAVAGGGGRMRFSVSSPDE 203
Db 122 ETMQKAGVTPVPGSEGLVETEQEGLE----LAKDIGYPMVNIKATAGGGGRMLVRSPPDE 177

QY 204 LRKLATEASREAAFGDGSVYVERAVINPQIHIEVOILGDRGTGEVVLHYERDCSLQRRHQ 263
Db 178 FVKFLAAGGAGAAFGNAGVYIEKPIERPHIEFOILADNYGNVTHLGERDCSLQRRNQ 237

QY 264 KVVETAPAGHLPDELDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEMNPRIOVE 323
Db 238 KLLLEAPSPALSDUREKMGQAARAAQFINTYAGTIEBFLDORSQGFYFEMNTRIOVE 297

```

```

QY 324 HTVTEEVTEVDLVKAOMRLAAGATLKEGLTQDKIKTHGAALOCRTTTPNNGFRPDTG 383
Db 298 HPVTMVTGVDLLVEQIRIAOGERLR---LTQDQVVLRGHAIIECRINAEDPDHDFRPA 354

QY 384 TITAYRSPGGAGVRLDGAAGLGEITAHFDSMLVKMTCRGSDFETAVARAQALAEFTVS 443
Db 355 RISGYLPPGGVGRIDSHVYTDYQIPYVDSLIGLVKLVMPDRATARNMKRALRECAIT 414

QY 444 GYATNIGFLRALLREEDFTYSKRIATGFI 471
Db 415 GLPTTIGFHQRIMENPQFGNVSTSFV 442

RESULT 7
US-08-422-560A-6
; Sequence 6, Application US/08422560A
; Patent No. 5910626
; GENERAL INFORMATION:
; APPLICANT: Haselkorn, Robert
; APPLICANT: Gornicki, Piotr
; TITLE OF INVENTION: ACETYL-CoA CARBOXYLASE COMPOSITIONS AND
; TITLE OF INVENTION: METHODS FOR USE
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/422,560A
; FILING DATE: 14-APR-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/956,700
; FILING DATE: 02-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Wilson, Mark B.
; REGISTRATION NUMBER: 37,259
; REFERENCE/DOCKET NUMBER: ARCD:152/WIM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512-418-3000
; TELEFAX: 512-474-7577
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 447 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
;
US-08-422-560A-6
Query Match 17.6%; Score 1032.5; DB 2; Length 447;
Best Local Similarity 48.2%; Pred. No. 1.4e-73;
Matches 216; Conservative 71; Mismatches 148; Indels 13; Gaps 4;

QY 29 FKKILVANGETIAVRAALETGAATVAIYPREDGSGFHRSEAVRIGTGSPVKAY 88
Db 3 FDKILLANGETIALRLTRACEEMGIATIAVSHVTDNRNALHVQLADEAVCGIG-EPASAKSY 61

QY 89 LDIDETIGAAGKVKADAIYPGYGFLSENALARECAENGITFTGPTPEVLDLTGDKSRV 148
Db 62 LNIPIIAAALTRNASHAIHPGYGFLSENALARECAENGITFTGPTPEVLDLTGDKSRV 121

QY 149 TAAKAGLPLVAES-----TPSKNIDDIVKSAEGQTYPIFVKAVAGGGGRMRFSVSSPDE 203
Db 122 ETMQKAGVTPVPGSEGLVETEQEGLE----LAKDIGYPMVNIKATAGGGGRMLVRSPPDE 177

```

QY 204 LRKLTASREAEAAFGDGSVYVERAVINPOHIEVQILGDRGTGEVHLXERDCSLORRHQ 263
DB 178 FVKFLAAQGEAGAAFGAGVYIEKFTERPRIEFQILADNYGNVHLGERDCSIORRNQ 237
QY 264 KVVETAPQHLDPRLDRICADAVKFCRSIGYOCAGTVEFLVDEKGNHVFIEHNPRIQVE 323
DB 238 KLEEAAPSPALDSLRKMGQAAVKAQAQFINYTGAQTIFFLLDRSGQFYFEMNTRIQVE 297
QY 324 HTVTEEVTEVDLVKAQMLAAGATLKLGLTODKIKTHGAALOCRTTDPNNGFRPDTG 383
DB 298 HPVTMTGVDLLVEQIRIAQGERLR---LTQDQVVLGRHAEICRINAEDPDHDFRAPG 354
QY 384 TITAYRSPGAGVRLDGAALGGEITAFHDSMLVKWTCRSDDETAVARAQALAEFTVS 443
DB 355 RISGLPPGGPGVRIDSHVYTDQIPYDLSLIGKLVWGPDRATINRMKRALRECAIT 414
QY 444 GVATNIGFLRALLREEDFTSKRIATGFI 471
DB 415 GLPTTIGFHQRIMENPQFLQGNVSTSFV 442

RESULT 8
US-08-468-793-6
; Sequence 6, Application US/08468793
; Patent No. 6177267
; GENERAL INFORMATION:
; APPLICANT: Haselkorn, Robert
; APPLICANT: Gornicki, Piotr
; TITLE OF INVENTION: ACETYL-COA CARBOXYLASE COMPOSITIONS AND
; TITLE OF INVENTION: METHODS OF USE
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,793
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/422,560
; FILING DATE: 14-APR-1995
; APPLICATION NUMBER: US SN 07/956,700
; FILING DATE: 02-OCT-1992
; CLASSIFICATION: 800
; APPLICATION NUMBER: PCT/US93/09340
; FILING DATE: 30-SEP-1993
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: ARCD:152/KIT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 447 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-468-793-6

Query Match 17.6%; Score 1032.5; DB 4; Length 447;
Best Local Similarity 48.2%; Pred. No. 1.4e-73;
Matches 216; Conservative 71; Mismatches 148; Indels 13; Gaps 4;
QY 29 FKKILVANGETAVAFRAALETGAATVAIYIPREDRGSFHRFSFASEAVRIGTEGSPVKAY 88
DB 3 FDKILIANRGEIALRLRACEEMGIATIAVHSTVDNRNALHVGLADEAVCIG-EPASAKSY 61
QY 89 LDIDEITGAAKVKADATYPGVGFSENAQALARECAENGITFIPPTPEVLDLTGDKSRV 148
DB 62 LNPINIAAALTRNSAHPGCGFLENAKFAEICADHIAFIGTTPAIRLMGDKSTAK 121
QY 149 TAAKAGLPLVAES-----TPSKNIDDIVKSAGQTYPIFVAVAGGGGRMRFFVSSPDE 203
DB 122 ETMQKAGVTPVPGSEGLVETEGL----LAKDIGPVMKATAGGGGGRMLRVSRPDE 177
QY 204 LRKLTASREAEAAFGDGSVYVERAVINPOHIEVQILGDRGTGEVHLXERDCSLORRHQ 263
DB 178 FVKFLAAQGEAGAAFGAGVYIEKFTERPRIEFQILADNYGNVHLGERDCSIORRNQ 237
QY 264 KVVETAPQHLDPRLDRICADAVKFCRSIGYOCAGTVEFLVDEKGNHVFIEHNPRIQVE 323
DB 238 KLEEAAPSPALDSLRKMGQAAVKAQAQFINYTGAQTIFFLLDRSGQFYFEMNTRIQVE 297
QY 324 HTVTEEVTEVDLVKAQMLAAGATLKLGLTODKIKTHGAALOCRTTDPNNGFRPDTG 383
DB 298 HPVTMTGVDLLVEQIRIAQGERLR---LTQDQVVLGRHAEICRINAEDPDHDFRAPG 354
QY 384 TITAYRSPGAGVRLDGAALGGEITAFHDSMLVKWTCRSDDETAVARAQALAEFTVS 443
DB 355 RISGLPPGGPGVRIDSHVYTDQIPYDLSLIGKLVWGPDRATINRMKRALRECAIT 414
QY 444 GVATNIGFLRALLREEDFTSKRIATGFI 471
DB 415 GLPTTIGFHQRIMENPQFLQGNVSTSFV 442

RESULT 9
US-08-611-107-8
; Sequence 8, Application US/08611107
; Patent No. 5801233
; GENERAL INFORMATION:
; APPLICANT: Haselkorn, Robert
; APPLICANT: Gornicki, Piotr
; TITLE OF INVENTION: NUCLEIC ACID COMPOSITIONS ENCODING
; TITLE OF INVENTION: ACETYL-COA CARBOXYLASE AND USES
; TITLE OF INVENTION: THEREFOR
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/611,107
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US SN 07/956,700
; FILING DATE: 02-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US93/09340
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US SN 08/422,560

```
; FILING DATE: 14-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: ARCD:221
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 453 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-08-611-107-8

Query Match 17.1%; Score 1005.5; DB 1; Length 453;
Best Local Similarity 47.1%; Pred. No. 2e-71;
Matches 209; Conservative 72; Mismatches 158; Indels 5; Gaps 3;

Qy 29 FKKILVANRGEIATVAFRAALETGAATVAIYPREDRGSFHRSEAFSEAVRIGTEGSPVKAY 88
Db 3 FKKILIANRGEIATVAFRAALETGAATVAIYPREDRGSFHRSEAFSEAVRIGTEGSPVKAY 88
Qy 89 LDIDEITGAAKKVKADAIYPGYFLSENAQLARECAENGITFTGPTPEVLDLTGDKSRV 148
Db 62 LNIPIIAAALTRNASAIHPGYGFLAENARFAEICADHHLTFIGSPDSIRAMGDKSTAK 121
Qy 149 TAAKAGLPVLAESTP-SKNIDDIIVKSAEGOTYPIFYKAVAGGGGRMRFVSSPDELRLK 207
Db 122 ETMRGVGPTIPGSDGLLTVDVSAKVAEEIGYPMVKATAGGGGRMRLVREPADLEKL 181
Qy 208 ATEASRAEAAFGDGSVYVERAVINPOHIEVQILGDRTEGVVHLYERDCSLQRHOKVVE 267
Db 182 FLAAGGEAAAFGNPGLYLEKFDPRHVEFQILADAGNVVHGERDCSIQRHOKLLE 241
Qy 268 IAPAQHLDPELRDICAADAVKFCRSIGYGAGTVEFLVDEKGNHVFIEMNPRIOVEHTVT 327
Db 242 EAPSPALSADLRQKMGDAAVKVAQAIYIGAGTVEFLVDATGNTFYFEMMNTRIQVEHPT 301
Qy 328 EEVTEVDLVKAQMLAAGATLKEGLTQDKIKTHGAALQCRITTEDPNNGRFPDGTGTTA 387
Db 302 EMITGLDLIAEQIRIAQGEALR---FROADIQLRGHATECRINAEDPEYFNRPNGRITG 358
Qy 388 YRSPGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDFETAVARAQALAEFTVSGVAT 447
Db 359 YLPPGGVGRVDSHVYTDYEIPPPYDLSLIGKLIWNGATREEARIMQALRECAITGLPT 418
Qy 448 NIGFLRALLREEDFTSKRIATGFI 471
Db 419 TLSFHQLMLQMPFELRGELYTNFV 442

RESULT 10
US-08-422-560A-8
; Sequence 8, Application US/08422560A
; Patent No. 5910626
; GENERAL INFORMATION:
; APPLICANT: Haselkorn, Robert
; APPLICANT: Gornicki, Piotr
; TITLE OF INVENTION: ACETYL-CoA CARBOXYLASE COMPOSITIONS AND
; TITLE OF INVENTION: METHODS FOR USE
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/422,560A
; FILING DATE: 14-APR-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/956,700
; FILING DATE: 02-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Wilson, Mark B.
; REGISTRATION NUMBER: 37,259
; REFERENCE/DOCKET NUMBER: ARCD:152/WIM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512-418-3000
; TELEFAX: 512-474-7577
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 453 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-08-422-560A-8

Query Match 17.1%; Score 1005.5; DB 2; Length 453;
Best Local Similarity 47.1%; Pred. No. 2e-71;
Matches 209; Conservative 72; Mismatches 158; Indels 5; Gaps 3;

Qy 29 FKKILVANRGEIATVAFRAALETGAATVAIYPREDRGSFHRSEAFSEAVRIGTEGSPVKAY 88
Db 3 FKKILIANRGEIATVAFRAALETGAATVAIYPREDRGSFHRSEAFSEAVRIGTEGSPVKAY 88
Qy 89 LDIDEITGAAKKVKADAIYPGYFLSENAQLARECAENGITFTGPTPEVLDLTGDKSRV 148
Db 62 LNIPIIAAALTRNASAIHPGYGFLAENARFAEICADHHLTFIGSPDSIRAMGDKSTAK 121
Qy 149 TAAKAGLPVLAESTP-SKNIDDIIVKSAEGOTYPIFYKAVAGGGGRMRFVSSPDELRLK 207
Db 122 ETMRGVGPTIPGSDGLLTVDVSAKVAEEIGYPMVKATAGGGGRMRLVREPADLEKL 181
Qy 208 ATEASRAEAAFGDGSVYVERAVINPOHIEVQILGDRTEGVVHLYERDCSLQRHOKVVE 267
Db 182 FLAAGGEAAAFGNPGLYLEKFDPRHVEFQILADAGNVVHGERDCSIQRHOKLLE 241
Qy 268 IAPAQHLDPELRDICAADAVKFCRSIGYGAGTVEFLVDEKGNHVFIEMNPRIOVEHTVT 327
Db 242 EAPSPALSADLRQKMGDAAVKVAQAIYIGAGTVEFLVDATGNTFYFEMMNTRIQVEHPT 301
Qy 328 EEVTEVDLVKAQMLAAGATLKEGLTQDKIKTHGAALQCRITTEDPNNGRFPDGTGTTA 387
Db 302 EMITGLDLIAEQIRIAQGEALR---FROADIQLRGHATECRINAEDPEYFNRPNGRITG 358
Qy 388 YRSPGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDFETAVARAQALAEFTVSGVAT 447
Db 359 YLPPGGVGRVDSHVYTDYEIPPPYDLSLIGKLIWNGATREEARIMQALRECAITGLPT 418
Qy 448 NIGFLRALLREEDFTSKRIATGFI 471
Db 419 TLSFHQLMLQMPFELRGELYTNFV 442

RESULT 11
US-08-468-793-8
; Sequence 8, Application US/08468793
; Patent No. 6177267
; GENERAL INFORMATION:
; APPLICANT: Haselkorn, Robert
; APPLICANT: Gornicki, Piotr
; TITLE OF INVENTION: ACETYL-CoA CARBOXYLASE COMPOSITIONS AND
; TITLE OF INVENTION: METHODS OF USE
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
```

STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,793
FILING DATE: 06-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/422,560
FILING DATE: 14-APR-1995
APPLICATION NUMBER: US SN 07/956,700
FILING DATE: 02-OCT-1992
CLASSIFICATION: 800
APPLICATION NUMBER: PCT/US93/09340
FILING DATE: 30-SEP-1993
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: ARCD:152/KIT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 453 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-468-793-8

Query Match 17.1%; Score 1005.5; DB 4; Length 453;
Best Local Similarity 47.1%; Pred. No. 2e-71;
Matches 209; Conservative 72; Mismatches 158; Indels 5; Gaps 3;
QY 29 FKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSPHRSFASAVRIGTEGSPVKAY 88
DB 3 FNKILIANRGEIALRILRTCEELGIGTIAVHSTVDNRNALHVQLADEAVCIG-EAASSKSY 61
QY 89 LDIDEITGAARKVKADAIYPGYGFLSNAQALARECAENGITFTIGPTPEVLDLTGDKSRV 148
DB 62 LNIPIIAAALTRNASAIHPGYGFLAENARFAEICADHHLTFIGSPDSIRAMGDKSTAK 121
QY 149 TAAKAGLPLVAESTP-SKNIDDIVKSAEGQYPIFVKAVAGGGRGMRVSSPDELRLK 207
DB 122 ETMQRVGPTIPGSDGLLTVDVDSAAKVAEIGYPMIKATAGGGGMRMLVREPADLEKL 181
QY 208 ATEASRAEAAFGDGSVYVERAVINPOHIEVQILGDRTEGVVHLYERDCSLORRHQKVE 267
DB 182 FLAAQGEAAEAFGNPGLYLEKIDRPHVEGQILADAYGNVHVGDERCSIORRHQKLE 241
QY 268 IAPAQHLDELDRICADAVKFCRSIGYAGTVEFLVDEKGNHVFEMNPRIOVHTVT 327
DB 242 EAPSPALSADLRQKMGDAAVKVAQAIYIGAGTVEFLVDATGNFYFEMNTRIQQVHPVT 301
QY 328 EEVTEVDLVKAQMLAAGATLKEGLTQDKIKTHGAALOCRIITTEDPNNGFRPDTGTITA 387
DB 302 EMITGLDLIAEQIRIAQGEALR---FRQADIQLRGAIECRINAEDPEYFNPNPGRITG 358
QY 388 YRSPGGAGVRLDGAALGGGEITAHFDSMLVKMTCRGSDFETAVARAQALAEFTVSGVAT 447
DB 359 YLPPGGPGVRVDSHVYTDYEIPPYDSLIGKLIWVGATREAIARMQALRECAITGLPT 418
QY 448 NIGFLRALLREEDFTSKRIATGFI 471

RESULT 13
US-07-956-700B-6
Sequence 6, Application US/07956700B
Patent No. 5539092
GENERAL INFORMATION:
APPLICANT: Robert Haselkorn and Piotr Gornicki

DB 419 TLSFHQLMLQMPFELRGELYTNFV 442
RESULT 12
US-09-433-043B-121
Sequence 121, Application US/09433043B
Patent No. 6399342
GENERAL INFORMATION:
APPLICANT: HASSELKORN, ROBERT
APPLICANT: GORNICKI, PIOTR
TITLE OF INVENTION: CYANOBACTERIAL AND PLANT ACETYL-CoA CARBOXYLASE
FILE REFERENCE: ARCD:338US
CURRENT APPLICATION NUMBER: US/09/433,043B
CURRENT FILING DATE: 1999-10-25
PRIOR FILING DATE: 08/475,879
PRIOR APPLICATION NUMBER: 07/956,700
PRIOR FILING DATE: 1992-10-02
NUMBER OF SEQ ID NOS: 128
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 121
LENGTH: 453
TYPE: PPT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: Peptide
US-09-433-043B-121

Query Match 17.1%; Score 1003.5; DB 4; Length 453;
Best Local Similarity 47.1%; Pred. No. 2.9e-71;
Matches 209; Conservative 72; Mismatches 158; Indels 5; Gaps 3;
QY 29 FKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSPHRSFASAVRIGTEGSPVKAY 88
DB 3 FNKILIANRGEIALRILRTCEELGIGTIAVHSTVDNRNALHVQLADEAVCIG-EAASSKSY 61
QY 89 LDIDEITGAARKVKADAIYPGYGFLSNAQALARECAENGITFTIGPTPEVLDLTGDKSRV 148
DB 62 LNIPIIAAALTRNASAIHPGYGFLAENARFAEICADHHLTFIGSPDSIRAMGDKSTAK 121
QY 149 TAAKAGLPLVAESTP-SKNIDDIVKSAEGQYPIFVKAVAGGGRGMRVSSPDELRLK 207
DB 122 ETMQRVGPTIPGSDGLLTVDVDSAAKVAEIGYPMIKATAGGGGMRMLVREPADLEKL 181
QY 208 ATEASRAEAAFGDGSVYVERAVINPOHIEVQILGDRTEGVVHLYERDCSLORRHQKVE 267
DB 182 FLAAQGEAAEAFGNPGLYLEKIDRPHVEGQILADAYGNVHVGDERCSIORRHQKLE 241
QY 268 IAPAQHLDELDRICADAVKFCRSIGYAGTVEFLVDEKGNHVFEMNPRIOVHTVT 327
DB 242 EAPSPALSADLRQKMGDAAVKVAQAIYIGAGTVEFLVDATGNFYFEMNTRIQQVHPVT 301
QY 328 EEVTEVDLVKAQMLAAGATLKEGLTQDKIKTHGAALOCRIITTEDPNNGFRPDTGTITA 387
DB 302 EMITGLDLIAEQIRIAQGEALR---FRQADIQLRGAIECRINAEDPEYFNPNPGRITG 358
QY 388 YRSPGGAGVRLDGAALGGGEITAHFDSMLVKMTCRGSDFETAVARAQALAEFTVSGVAT 447
DB 359 YLPPGGPGVRVDSHVYTDYEIPPYDSLIGKLIWVGATREAIARMQALRECAITGLPT 418
QY 448 NIGFLRALLREEDFTSKRIATGFI 471
DB 419 TLSFHQLMLQMPFELRGELYTNFV 442

RESULT 13
US-07-956-700B-6
Sequence 6, Application US/07956700B
Patent No. 5539092
GENERAL INFORMATION:
APPLICANT: Robert Haselkorn and Piotr Gornicki

: TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA
: TITLE OF INVENTION: Carboxylase
: NUMBER OF SEQUENCES: 116
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Arnold, White & Durkee
: STREET: 321 No. 553092th Clark Street
: CITY: Chicago
: STATE: Illinois
: COUNTRY: USA
: ZIP: 60610
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC Compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: ASCII-DOS
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/956,700B
: FILING DATE: 19921002
: CLASSIFICATION: 800
: ATTORNEY/AGENT INFORMATION:
: NAME: Thomas E. No. 553092thrup
: REGISTRATION NUMBER: 33,268
: REFERENCE/DOCKET NUMBER: ARCD:058
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 1-312-744-0090
: TELEFAX: 1-312-755-4489
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 453 amino acids
: TYPE: Amino acid
: STRANDEDNESS: Single
: TOPOLOGY: Linear
: MOLECULE TYPE: Peptide
: US-07-956-700B-6

Query Match 17.0%; Score 996.5; DB 1; Length 453;
Best Local Similarity 46.8%; Pred. No. 1e-70;
Matches 208; Conservative 72; Mismatches 159; Indels 5; Gaps 3;

QY 29 FKKILVANRGEIAVRAFAALETGAATVAIYPREDRSGFHRSEAFSEAVRIGTEGSPVKAY 88
Db 3 FNKILIANRGEIALRILRCEELGIGTIAVHSTVDNRNALHVQLADEAVCIG-EAASSKSY 61
QY 89 LDIDEIIGAIAKKVADAIYPGYGFLSENAQLARECAENGITFGTPEVLDLTGDKSRV 148
Db 62 LNPINIIAALTRNSAHPGYGFLAENARFAEICADHHLTFTIGSPDSIRAMGDKSTAK 121
QY 149 TAAKKAGLPVLAESTP-SKNIDDIVKSAEGQTPYIFVKAVAGGGGRMRVSSPDELRLK 207
Db 122 ETMQRVGVPTIPGSDGLLTVDVSAAKVAEEIGYPMVMIKATAGGGGRMRLVREPADLEK 181
QY 208 ATEASREAEAAFGDGSVYVERAVINPOHIEVOILGDRTEGVVHLYERDCSLORRHOKVVE 267
Db 182 FLAQQGEAEAAFGNPGLYLEKFDPRPHVEFQILADAYGNVVELGERDCSIQRHOKLLE 241
QY 268 IAPQHLDPRLDRICADAVKFCRSICYGAGTVEFLVDEKGNHVFTEMPNRIQVEHTVT 327
Db 242 EAPSPALSADLRQKMGDAAVKVAQAIGYICAGTVEFLVDATGNFYFNMNTRIQLVEHPVT 301
QY 328 EEVTEVDLVKAQMRLAAGATLKELGILTQDKIKTHGAALQCRITTEDPNNGFRPDTGITTA 387
Db 302 EMITGLDLIAEQIRIAQGEALR---FRQADIQLRGHAEICRNAEDPEYFNPNPGRITG 358
QY 388 YRSPGCGAVRLDGAQAALGCGEITAHFDSMLVKMTCRGSDPETAVARAQALAEFTVSGVAT 447
Db 359 YLPDPGGVGRVDSHVYTDYIPYDLSLIGKLIWGTATREEARQALREGAITGLPT 418
QY 448 NIGFLLALLREEDFTSKRIATGFI 471
Db 419 TLSFHQLMLQMPPEFLRGELYTNEV 442

RESULT 14

US-08-476-537-6
: Sequence 6, Application US/08476537
: Patent No. 5756290
: GENERAL INFORMATION:
: APPLICANT: Robert Haselkorn and Piotr Gornicki
: TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA
: NUMBER OF SEQUENCES: 116
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Arnold, White & Durkee
: STREET: 321 No. 5756290th Clark Street
: CITY: Chicago
: STATE: Illinois
: COUNTRY: USA
: ZIP: 60610
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC Compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: ASCII-DOS
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/476,537
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/956,700
: FILING DATE: 10/21/92
: ATTORNEY/AGENT INFORMATION:
: NAME: Thomas E. No. 5756290thrup
: REGISTRATION NUMBER: 33,268
: REFERENCE/DOCKET NUMBER: ARCD:058
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 1-312-744-0090
: TELEFAX: 1-312-755-4489
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 453 amino acids
: TYPE: Amino acid
: STRANDEDNESS: Single
: TOPOLOGY: Linear
: MOLECULE TYPE: Peptide
: US-08-476-537-6

Query Match 17.0%; Score 996.5; DB 1; Length 453;
Best Local Similarity 46.8%; Pred. No. 1e-70;
Matches 208; Conservative 72; Mismatches 159; Indels 5; Gaps 3;

QY 29 FKKILVANRGEIAVRAFAALETGAATVAIYPREDRSGFHRSEAFSEAVRIGTEGSPVKAY 88
Db 3 FNKILIANRGEIALRILRCEELGIGTIAVHSTVDNRNALHVQLADEAVCIG-EAASSKSY 61
QY 89 LDIDEIIGAIAKKVADAIYPGYGFLSENAQLARECAENGITFGTPEVLDLTGDKSRV 148
Db 62 LNPINIIAALTRNSAHPGYGFLAENARFAEICADHHLTFTIGSPDSIRAMGDKSTAK 121
QY 149 TAAKKAGLPVLAESTP-SKNIDDIVKSAEGQTPYIFVKAVAGGGGRMRVSSPDELRLK 207
Db 122 ETMQRVGVPTIPGSDGLLTVDVSAAKVAEEIGYPMVMIKATAGGGGRMRLVREPADLEK 181
QY 208 ATEASREAEAAFGDGSVYVERAVINPOHIEVOILGDRTEGVVHLYERDCSLORRHOKVVE 267
Db 182 FLAQQGEAEAAFGNPGLYLEKFDPRPHVEFQILADAYGNVVELGERDCSIQRHOKLLE 241
QY 268 IAPQHLDPRLDRICADAVKFCRSICYGAGTVEFLVDEKGNHVFTEMPNRIQVEHTVT 327
Db 242 EAPSPALSADLRQKMGDAAVKVAQAIGYICAGTVEFLVDATGNFYFNMNTRIQLVEHPVT 301
QY 328 EEVTEVDLVKAQMRLAAGATLKELGILTQDKIKTHGAALQCRITTEDPNNGFRPDTGITTA 387
Db 302 EMITGLDLIAEQIRIAQGEALR---FRQADIQLRGHAEICRNAEDPEYFNPNPGRITG 358
QY 388 YRSPGCGAVRLDGAQAALGCGEITAHFDSMLVKMTCRGSDPETAVARAQALAEFTVSGVAT 447

GenCore version 5.1.4_p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2003, 19:02:46 ; Search time 24 Seconds
(without alignments)
2830.970 Million cell updates/sec

Title: US-09-974-973-2
Perfect score: 5865
Sequence: 1 MTAITLGGLLKGITITVST.....RVVPAATKVEGGDLIVVVS 1157

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 237916 seqs, 58723674 residues

Total number of hits satisfying chosen parameters: 237916

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5865	100.0	1157	9	US-09-974-973-2
2	5865	100.0	1157	9	US-09-974-973-4
3	5759	98.2	1140	9	US-09-974-973-19
4	5759	98.2	1140	9	US-09-738-626-4265
5	5759	98.2	1140	9	US-10-045-072-2
6	2489	42.4	1147	10	US-09-815-242-5468
7	2460.5	42.0	1142	10	US-09-815-242-10806
8	2300	39.2	1073	10	US-09-815-242-12361
9	1032.5	17.6	447	10	US-09-767-479-6
10	1005.5	17.1	453	10	US-09-767-479-8
11	979.5	16.7	471	10	US-09-815-242-5215
12	956	16.3	449	10	US-09-815-242-13885
13	950	16.2	448	10	US-09-815-242-11160
14	940	16.0	449	10	US-09-815-242-10330
15	936.5	16.0	455	10	US-09-815-242-13617
16	935.5	16.0	455	10	US-09-815-242-13364
17	921	15.7	449	10	US-09-815-242-12063
18	919	15.7	458	10	US-09-815-242-11321
19	909	15.5	443	10	US-09-815-242-4963

20	909	15.5	456	10	US-09-815-242-10924	Sequence 10324, A
21	902	15.4	455	10	US-09-815-242-11558	Sequence 11558, A
22	880.5	15.0	591	9	US-09-738-626-6940	Sequence 6940, Ap
23	856.5	14.6	446	10	US-09-815-242-5418	Sequence 5418, Ap
24	856.5	14.6	453	10	US-09-815-242-12562	Sequence 12562, A
25	837	14.3	448	10	US-09-815-242-5806	Sequence 5806, Ap
26	834	14.2	451	10	US-09-815-242-13127	Sequence 13127, A
27	633	10.8	358	10	US-09-815-242-12939	Sequence 12939, A
28	583.5	9.9	2257	10	US-09-767-479-10	Sequence 10, Appl
29	437.5	7.5	483	9	US-10-083-357-1328	Sequence 1328, Ap
30	377.5	6.4	158	10	US-09-815-242-5031	Sequence 5031, Ap
31	300.5	5.1	163	9	US-09-895-913A-324	Sequence 324, App
32	251	4.3	124	10	US-09-205-658-240	Sequence 240, App
33	244.5	4.2	1687	9	US-10-094-679-3	Sequence 3, Appli
34	225	3.8	124	10	US-09-205-658-238	Sequence 238, App
35	213.5	3.6	1113	9	US-09-738-626-5279	Sequence 5279, Ap
36	208.5	3.6	98	10	US-09-815-242-12938	Sequence 12938, A
37	208.5	3.6	1113	10	US-09-836-470B-3	Sequence 3, Appli
38	185	3.2	4999	9	US-09-976-059-15	Sequence 15, Appl
39	170	2.9	1289	9	US-09-712-363-259	Sequence 259, App
40	165.5	2.8	2969	9	US-09-738-626-4434	Sequence 4434, Ap
41	163.5	2.8	5215	9	US-09-860-846-2	Sequence 2, Appli
42	163.5	2.8	5215	10	US-09-861-289-2	Sequence 2, Appli
43	162.5	2.8	122	10	US-09-987-485-1	Sequence 1, Appli
44	160	2.7	4999	9	US-09-976-059-14	Sequence 14, Appl
45	158.5	2.7	70	10	US-09-987-485-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-974-973-2
; Sequence 2, Application US/09974973
; Patent No. US20020177202A1
; GENERAL INFORMATION:
; APPLICANT: Hanke, Paul D.
; TITLE OF INVENTION: Feedback-Resistant Pyruvate Carboxylase Gene from Corynebacter
; FILE REFERENCE: 1533.1230001/MAC/RGM
; CURRENT APPLICATION NUMBER: US/09/974,973
; CURRENT FILING DATE: 2001-10-21
; PRIOR APPLICATION NUMBER: US 60/239,913
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1157
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-974-973-2

Query Match	100.0%	Score 5865;	DB 9;	Length 1157;
Best Local Similarity	100.0%	Pred. No. 0;	Mismatches 0;	Indels 0;
Matches 1157;	Conservative	0;		Gaps 0;
QY	1	MTAITLGGLLKGITITVSTHTSSTLPAPFKKILVANRGEIAVRAALETGAATVAIYP	60	
Db	1	MTAITLGGLLKGITITVSTHTSSTLPAPFKKILVANRGEIAVRAALETGAATVAIYP	60	
QY	61	REDRGSFHRSFASEAVRICTEGSPVKAYLDIDEIIGAANKVKADATIPGYGFLSENAQLA	120	
Db	61	REDRGSFHRSFASEAVRICTEGSPVKAYLDIDEIIGAANKVKADATIPGYGFLSENAQLA	120	
QY	121	RECAENGITFIPTPEVLDTGDKSRVTAANKAGLPLVLAESTPSKNIDDIIVKSAEQTY	180	
Db	121	RECAENGITFIPTPEVLDTGDKSRVTAANKAGLPLVLAESTPSKNIDDIIVKSAEQTY	180	
QY	181	PIFVKAVAGGGGMRFSVSPDELRLKLA TEASREAAAFGDSGVYVERAVINPOHLEVOI	240	
Db	181	PIFVKAVAGGGGMRFSVSPDELRLKLA TEASREAAAFGDSGVYVERAVINPOHLEVOI	240	
QY	241	LGDRTEGVVHLRYDCSLQRRHQKVVEIAPAQHLDELRLDRICADAVKFCRSIGYOGAGT	300	
Db	241	LGDRTEGVVHLRYDCSLQRRHQKVVEIAPAQHLDELRLDRICADAVKFCRSIGYOGAGT	300	

```
Db 241 LGDRTGEVHVHLYERDCSLQRRHQKVVETIAPAQHLDPELDRICADAVKFCRSIGYQAGT 300
Qy 301 VEFVLDEKGNHVFIEIMPRIQVEHTVEEVDLVKAQMRLAAGATLKEIGLGTQDKIKT 360
Db 301 VEFVLDEKGNHVFIEIMPRIQVEHTVEEVDLVKAQMRLAAGATLKEIGLGTQDKIKT 360
Qy 361 HGAALQCRITTTEDPNNCFRDPDTGTTAYRSPGGAGVRLDGAALQGGEITAHFDSMLVKMT 420
Db 361 HGAALQCRITTTEDPNNCFRDPDTGTTAYRSPGGAGVRLDGAALQGGEITAHFDSMLVKMT 420
Qy 421 CRGSDFFETAVARAQALAEFTVSGVATNIGFLRALLREEDFTSKRIATGTFIDGPHLLQA 480
Db 421 CRGSDFFETAVARAQALAEFTVSGVATNIGFLRALLREEDFTSKRIATGTFIDGPHLLQA 480
Qy 481 PPADDEQGRILDLADVTYVKNKPHGVKDVAAADPKLPNIKOLPLPRGSDRLKQLGPAA 540
Db 481 PPADDEQGRILDLADVTYVKNKPHGVKDVAAADPKLPNIKOLPLPRGSDRLKQLGPAA 540
Qy 541 FARDLREQDALAVTDTTFRDAHQSLATRVRSFALKPAAEAVAKLTPELLSVEAWGATY 600
Db 541 FARDLREQDALAVTDTTFRDAHQSLATRVRSFALKPAAEAVAKLTPELLSVEAWGATY 600
Qy 601 DVAMRFLFEDPMDRLDELREAMPNWNOMLLRGNTVGYTYPDSVCRAFVKEAASSGVD 660
Db 601 DVAMRFLFEDPMDRLDELREAMPNWNOMLLRGNTVGYTYPDSVCRAFVKEAASSGVD 660
Qy 661 IFRIFDALNDVSQMRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYILKMAEEIV 720
Db 661 IFRIFDALNDVSQMRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYILKMAEEIV 720
Qy 721 KSGAHILAIDKDMAGLLRPAAVTKLVTLALRRREFDLPVHVHTDTAGGOLATYFAAAQAGAD 780
Db 721 KSGAHILAIDKDMAGLLRPAAVTKLVTLALRRREFDLPVHVHTDTAGGOLATYFAAAQAGAD 780
Qy 781 AVDGASAPLSTTTSQPSLSAIVAAFAHTRDRTGLSLEAVSDLEPYWEAVRGLYLPFESGT 840
Db 781 AVDGASAPLSTTTSQPSLSAIVAAFAHTRDRTGLSLEAVSDLEPYWEAVRGLYLPFESGT 840
Qy 841 PGPTGRVYRHEIPGQSLNRAQATLGLADREFLIEDNYAANVEMLGRPTKVTTPSSKVV 900
Db 841 PGPTGRVYRHEIPGQSLNRAQATLGLADREFLIEDNYAANVEMLGRPTKVTTPSSKVV 900
Qy 901 GDALHLVGAGVDPADFAADPKYDIIPDSVIAFLRGELGNPPGWPPEPLTRALEGRSEG 960
Db 901 GDALHLVGAGVDPADFAADPKYDIIPDSVIAFLRGELGNPPGWPPEPLTRALEGRSEG 960
Qy 961 KAPLTEVPEEQAHLDADDSKERNNSLNRLLPKPTPEEFLEHRRRFGNTSALDDREFFYG 1020
Db 961 KAPLTEVPEEQAHLDADDSKERNNSLNRLLPKPTPEEFLEHRRRFGNTSALDDREFFYG 1020
Qy 1021 LVEGRETLIRLPDVVRTPLLVRLDLAISEPDDKGRNVVYANGQIRPMRVRDRSVESVTAT 1080
Db 1021 LVEGRETLIRLPDVVRTPLLVRLDLAISEPDDKGRNVVYANGQIRPMRVRDRSVESVTAT 1080
Qy 1081 AEKADSSNKGHVAAFPAGVVTVTVAEGDEVKAGDAVAIIEMKMEATITASVDGKIERYV 1140
Db 1081 AEKADSSNKGHVAAFPAGVVTVTVAEGDEVKAGDAVAIIEMKMEATITASVDGKIERYV 1140
Qy 1141 VPAATKVEGGDLIVVVS 1157
Db 1141 VPAATKVEGGDLIVVVS 1157
```

RESULT 2

```
US-09-974-973-4
: Sequence 4, Application US/09974973
: Patent No. US2002017202A1
: GENERAL INFORMATION:
: APPLICANT: Hake, Paul D.
: TITLE OF INVENTION: Feedback-Resistant Pyruvate Carboxylase Gene from Corynebacterium
: FILE REFERENCE: 1533-1230001/MAC/RGM
: CURRENT APPLICATION NUMBER: US/09/974,973
: CURRENT FILING DATE: 2001-10-21
```

```
: PRIOR APPLICATION NUMBER: US 60/239,913
: PRIOR FILING DATE: 2000-10-13
: NUMBER OF SEQ ID NOS: 19
: SOFTWARE: Patentin version 3.0
: SEQ ID NO 4
: LENGTH: 1157
: TYPE: PRT
: ORGANISM: Corynebacterium glutamicum
: US-09-974-973-4
```

Query Match 100.0%; Score 5865; DB 9; Length 1157;

Best Local Similarity 100.0%; Pred. No. 0; Matches 1157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTAITLGGLLLGKIIITLVSTHTSSSTLPAPKKILVANRGEIIVRAFRAALETGAATVAIYVP 60

Db 1 MTAITLGGLLLGKIIITLVSTHTSSSTLPAPKKILVANRGEIIVRAFRAALETGAATVAIYVP 60

Qy 61 REDRGSFHRFSASEAVRIGTEGSPVKAYLDIDEIIGAARKVKADAIYPGYFLSENAQLA 120

Db 61 REDRGSFHRFSASEAVRIGTEGSPVKAYLDIDEIIGAARKVKADAIYPGYFLSENAQLA 120

Qy 121 RECAENGITFTIGTPEVLDLTGDKSRVTAAKKAGLPVLAESTPSKNIDDIIVKSAEGQTY 180

Db 121 RECAENGITFTIGTPEVLDLTGDKSRVTAAKKAGLPVLAESTPSKNIDDIIVKSAEGQTY 180

Qy 181 PIFVKAVAGGGGGRMRFVSSPDRLKLAETASREAAAFDGGSVYVERAVINPQHIEVOI 240

Db 181 PIFVKAVAGGGGGRMRFVSSPDRLKLAETASREAAAFDGGSVYVERAVINPQHIEVOI 240

Qy 241 LGDRTGEVHVHLYERDCSLQRRHQKVVETIAPAQHLDPELDRICADAVKFCRSIGYQAGT 300

Db 241 LGDRTGEVHVHLYERDCSLQRRHQKVVETIAPAQHLDPELDRICADAVKFCRSIGYQAGT 300

Qy 301 VEFVLDEKGNHVFIEIMPRIQVEHTVEEVDLVKAQMRLAAGATLKEIGLGTQDKIKT 360

Db 301 VEFVLDEKGNHVFIEIMPRIQVEHTVEEVDLVKAQMRLAAGATLKEIGLGTQDKIKT 360

Qy 361 HGAALQCRITTTEDPNNCFRDPDTGTTAYRSPGGAGVRLDGAALQGGEITAHFDSMLVKMT 420

Db 361 HGAALQCRITTTEDPNNCFRDPDTGTTAYRSPGGAGVRLDGAALQGGEITAHFDSMLVKMT 420

Qy 421 CRGSDFFETAVARAQALAEFTVSGVATNIGFLRALLREEDFTSKRIATGTFIDGPHLLQA 480

Db 421 CRGSDFFETAVARAQALAEFTVSGVATNIGFLRALLREEDFTSKRIATGTFIDGPHLLQA 480

Qy 481 PPADDEQGRILDLADVTYVKNKPHGVKDVAAADPKLPNIKOLPLPRGSDRLKQLGPAA 540

Db 481 PPADDEQGRILDLADVTYVKNKPHGVKDVAAADPKLPNIKOLPLPRGSDRLKQLGPAA 540

Qy 541 FARDLREQDALAVTDTTFRDAHQSLATRVRSFALKPAAEAVAKLTPELLSVEAWGATY 600

Db 541 FARDLREQDALAVTDTTFRDAHQSLATRVRSFALKPAAEAVAKLTPELLSVEAWGATY 600

Qy 601 DVAMRFLFEDPMDRLDELREAMPNWNOMLLRGNTVGYTYPDSVCRAFVKEAASSGVD 660

Db 601 DVAMRFLFEDPMDRLDELREAMPNWNOMLLRGNTVGYTYPDSVCRAFVKEAASSGVD 660

Qy 661 IFRIFDALNDVSQMRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYILKMAEEIV 720

Db 661 IFRIFDALNDVSQMRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYILKMAEEIV 720

Qy 721 KSGAHILAIDKDMAGLLRPAAVTKLVTLALRRREFDLPVHVHTDTAGGOLATYFAAAQAGAD 780

Db 721 KSGAHILAIDKDMAGLLRPAAVTKLVTLALRRREFDLPVHVHTDTAGGOLATYFAAAQAGAD 780

Qy 781 AVDGASAPLSTTTSQPSLSAIVAAFAHTRDRTGLSLEAVSDLEPYWEAVRGLYLPFESGT 840

Db 781 AVDGASAPLSTTTSQPSLSAIVAAFAHTRDRTGLSLEAVSDLEPYWEAVRGLYLPFESGT 840

Qy 841 PGPTGRVYRHEIPGQSLNRAQATLGLADREFLIEDNYAANVEMLGRPTKVTTPSSKVV 900

Db 841 PGPTGRVYRHEIPGQSLNRAQATLGLADREFLIEDNYAANVEMLGRPTKVTTPSSKVV 900

Qy	901	GDIALHLVGAGVDPADFAADPKYDIPDSVIAFLRGLGNPPGGWPEPLRTRALEGRSEG	960
Db	901	GDIALHLVGAGVDPADFAADPKYDIPDSVIAFLRGLGNPPGGWPEPLRTRALEGRSEG	960
Qy	961	KAPLTVPEEQAHILDADDSKERRNSLNRLFFPKPTEEFLEHRRRFGNTSALDDREFFYG	1020
Db	961	KAPLTVPEEQAHILDADDSKERRNSLNRLFFPKPTEEFLEHRRRFGNTSALDDREFFYG	1020
Qy	1021	LVGREGTLLRLPDVRTPLLVRLDAISEPDDKGMNVVANVNGOIRPMVRDRSRVESVTAT	1080
Db	1021	LVGREGTLLRLPDVRTPLLVRLDAISEPDDKGMNVVANVNGOIRPMVRDRSRVESVTAT	1080
Qy	1081	AEKADSSNKGHVAAPPAGVYVTVTAEGDEVKAGDAVAIEAMKMEATITASVDGKIERVV	1140
Db	1081	AEKADSSNKGHVAAPPAGVYVTVTAEGDEVKAGDAVAIEAMKMEATITASVDGKIERVV	1140
Qy	1141	VPAATKVEGGDLIVVVS	1157
Db	1141	VPAATKVEGGDLIVVVS	1157

RESULT 3
 US-09-974-973-19
 ; Sequence 19, Application US/09974973
 ; Patent No. US20020177202A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hanke, Paul D.
 ; TITLE OF INVENTION: Feedback-Resistant Pyruvate Carboxylase Gene from *Corynebacterium glutamicum*
 ; FILE REFERENCE: 1533.1230001/MAC/RGM
 ; CURRENT APPLICATION NUMBER: US/09/974,973
 ; CURRENT FILING DATE: 2001-10-21
 ; PRIOR APPLICATION NUMBER: US 60/239,913
 ; PRIOR FILING DATE: 2000-10-13
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 19
 ; LENGTH: 1140
 ; TYPE: PRT
 ; ORGANISM: *Corynebacterium glutamicum*
 US-09-974-973-19

Qy	438	A	EFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYADV	497
Db	421	A	EFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYADV	480
Qy	498	T	VNKPCHGVREPKDVAAPIDKLPNTKOLPLPRGSRDRLKQLGPAAPFARDLREQDALAVDTT	557
Db	481	T	VNKPCHGVREPKDVAAPIDKLPNTKOLPLPRGSRDRLKQLGPAAPFARDLREQDALAVDTT	540
Qy	558	F	RDHAQSHLLATRVRSFALPAAEAVAKLIPPELLSVKAEWGATYDVAMRFLFEDPWDRLDE	617
Db	541	F	RDHAQSHLLATRVRSFALPAAEAVAKLIPPELLSVKAEWGATYDVAMRFLFEDPWDRLDE	600
Qy	618	L	REAMPNVNTOMLLRGNTVGYTPYPDSVCRAFKVKEAASSGVDIFRIEDALNDVYSQMRPA	677
Db	601	L	REAMPNVNTOMLLRGNTVGYTPYPDSVCRAFKVKEAASSGVDIFRIEDALNDVYSQMRPA	660
Qy	678	I	DAVLENTAAVEANWAYSGDLSDPNEKLYTLDDYLYLKAAEETVKSAGHLLAIKDMAGLLR	737
Db	661	I	DAVLENTAAVEANWAYSGDLSDPNEKLYTLDDYLYLKAAEETVKSAGHLLAIKDMAGLLR	720
Qy	738	P	AAVTKLVTLARREFDPLPVHVHTHTDAGQLATYFAAAQAGADAVDGAAPLSGTTSQPS	797
Db	721	P	AAVTKLVTLARREFDPLPVHVHTHTDAGQLATYFAAAQAGADAVDGAAPLSGTTSQPS	780
Qy	798	L	SAIVAAFAHTRDGTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL	857
Db	781	L	SAIVAAFAHTRDGTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL	840
Qy	858	S	NLRAQATALGLADREFLETDNVAANVMEGLRPTKVTPSSKVVGDALHLVAGVDPAF	917
Db	841	S	NLRAQATALGLADREFLETDNVAANVMEGLRPTKVTPSSKVVGDALHLVAGVDPAF	900
Qy	918	A	ADPKQYDIPDSVIAFLRGLGNPPGGWPEPLTRALEGRSEKAPLTFVPEEQAHLDA	977
Db	901	A	ADPKQYDIPDSVIAFLRGLGNPPGGWPEPLTRALEGRSEKAPLTFVPEEQAHLDA	960
Qy	978	D	DSKERNLSNRLFLFKPTEEFLEHRRRGNTSALDDREFFYGLVEGRETLIRLPDVVRTP	1037
Db	961	D	DSKERNLSNRLFLFKPTEEFLEHRRRGNTSALDDREFFYGLVEGRETLIRLPDVVRTP	1020
Qy	1038	L	LVRLDAISEPDDKGMNVNNAVNGQIRPMRYRDRSVESVTATAEKADSSNKGHVAAPFA	1097
Db	1021	L	LVRLDAISEPDDKGMNVNNAVNGQIRPMRYRDRSVESVTATAEKADSSNKGHVAAPFA	1080
Qy	1098	G	VTVTVVAEGDEWKAGDAVAIIETAMKMEATITASVDGKIERYVVPAAATKVEGGDLIVVVS	1157
Db	1081	G	VTVTVVAEGDEWKAGDAVAIIETAMKMEATITASVDGKIDRVVVVPAATKVEGGDLIVVVS	1140
RESULT 4				
US-09-738-626-4265				
; Sequence 4265, Application US/09738626				
; Publication No. US20020197605A1				
; GENERAL INFORMATION:				
; APPLICANT: NAKAGAWA, SATOSHI				
; APPLICANT: MIZOGUCHI, HIROSHI				
; APPLICANT: ANDO, SEIKO				
; APPLICANT: HAYASHI, MIKIRO				
; APPLICANT: OCHIAI, KEIKO				
; APPLICANT: YOKOI, HARUHIKO				
; APPLICANT: TATEISHI, NAKO				
; APPLICANT: SENO, AKIHIRO				
; APPLICANT: IKEDA, NASATO				
; APPLICANT: OZAKI, AKIO				
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES				
; FILE REFERENCE: 249-125				
; CURRENT APPLICATION NUMBER: US/09/738,626				
; CURRENT FILING DATE: 2000-12-18				
; PRIOR APPLICATION NUMBER: JP 99/377484				
; PRIOR FILING DATE: 1999-12-16				
; PRIOR APPLICATION NUMBER: JP 00/159162				
; PRIOR FILING DATE: 2000-04-07				

; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4265
; LENGTH: 1140
; TYPE: PR1
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4265

Query Match 98.2%; Score 5759; DB 9; Length 1140;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1133; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 18 VSTHTSSTLPAPFKKILVANRGEIARAFRAALETGAATVAIYPREDRGSFHRSEAVR 77
Db 1 MSTHTSSTLPAPFKKILVANRGEIARAFRAALETGAATVAIYPREDRGSFHRSEAVR 60

QY 78 IGTEGSPVKAYLIDDEIIIGAARKKADAIYPGYFLSENAQLARECAENGITFTGPTPEV 137
Db 61 IGTEGSPVKAYLIDDEIIIGAARKKADAIYPGYFLSENAQLARECAENGITFTGPTPEV 120

QY 138 LDLTGDKSRANTAAKAGLPVLAESTPSKNIDDIIVKSAEGOTYPIFVKAVAGGGGRMRF 197
Db 121 LDLTGDKSRANTAAKAGLPVLAESTPSKNIDDIIVKSAEGOTYPIFVKAVAGGGGRMRF 180

QY 198 YSSPDELRLKLAESREAEAFGDSVYVERAVINPQIHIEVQIILGDRGTGEVHHLYERDCS 257
Db 181 VASPDDELRLKLAESREAEAFGDSVYVERAVINPQIHIEVQIILGDRGTGEVHHLYERDCS 240

QY 258 LQRRHQKVEIAPAOHLDPDLRDRICADAVKFCRSIGYQAGTGVFEVLVDEKGNHVFIEMN 317
Db 241 LQRRHQKVEIAPAOHLDPDLRDRICADAVKFCRSIGYQAGTGVFEVLVDEKGNHVFIEMN 300

QY 318 PRIQVEHTVTEEVTDVLVKAQMLAAGATLKLGLTQDKIKTHGAALQCRITTEDPNNG 377
Db 301 PRIQVEHTVTEEVTDVLVKAQMLAAGATLKLGLTQDKIKTHGAALQCRITTEDPNNG 360

QY 378 FRPDGTGITAYRSPGAGVRLDGAOLGGEITAHFDSMLVKMTCRGSDFETAVARAQAL 437
Db 361 FRPDGTGITAYRSPGAGVRLDGAOLGGEITAHFDSMLVKMTCRGSDFETAVARAQAL 420

QY 438 AEFTVSGVATNIGFRLALLREEDFTSKRIATGFIADPHLLQAPADDEOGRILLDYADY 497
Db 421 AEFTVSGVATNIGFRLALLREEDFTSKRIATGFIADPHLLQAPADDEOGRILLDYADY 480

QY 498 TVNKPBGVVRPKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAAPFARDLREODALAVTDTT 557
Db 481 TVNKPBGVVRPKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAAPFARDLREODALAVTDTT 540

QY 538 FRDAHQSLLATRVRSFALKPAEAAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWRDLDE 617
Db 541 FRDAHQSLLATRVRSFALKPAEAAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWRDLDE 600

QY 618 LREAMPNVIQMLLRGNTVGYTPYPSVCRAFVKEAASSGVDFIRIFDALNDVYSQMRPA 677
Db 601 LREAMPNVIQMLLRGNTVGYTPYPSVCRAFVKEAASSGVDFIRIFDALNDVYSQMRPA 660

QY 678 IDAVLETTNVAEAVMAYSGDLSDPNEKLYTLDYLLKMAEIVKSGAHILAIKDMAGLLR 737
Db 661 IDAVLETTNVAEAVMAYSGDLSDPNEKLYTLDYLLKMAEIVKSGAHILAIKDMAGLLR 720

QY 738 PAANTKLVTALRRFEDLPVHHVTHDTAGGQLATYFAAAQAGADAVDGCASAPLSTGTSQPS 797
Db 721 PAANTKLVTALRRFEDLPVHHVTHDTAGGQLATYFAAAQAGADAVDGCASAPLSTGTSQPS 780

QY 798 LSAIVAAFAHTRDTGLSLEAVSDELPWEAVRGLYLPFESGTPGPTGRVYRIEIPGGQL 857
Db 781 LSAIVAAFAHTRDTGLSLEAVSDELPWEAVRGLYLPFESGTPGPTGRVYRIEIPGGQL 840

QY 858 SNLRAQATFALGLADRFELIEDNYAAVNEMLRGPTKVTPTSSKVVGVGDALHLVAGVDPADF 917
Db 841 SNLRAQATFALGLADRFELIEDNYAAVNEMLRGPTKVTPTSSKVVGVGDALHLVAGVDPADF 900

RESULT 5
US-10-045-072-2
; Sequence 2, Application US/10045072
; Publication No. US20030027305A1
; GENERAL INFORMATION:
; APPLICANT: Sinskey, Anthony J.
; APPLICANT: Lessard, Philip A.
; APPLICANT: Willis, Laura B.
; TITLE OF INVENTION: Pyruvate Carboxylase from Corynebacterium glutamicum
; FILE REFERENCE: 1533.0790002
; CURRENT APPLICATION NUMBER: US/10/045,072
; CURRENT FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 09/677,575
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: US 09/220,081
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 2
; LENGTH: 1140
; TYPE: PR1
; ORGANISM: Corynebacterium glutamicum
US-10-045-072-2

Query Match 98.2%; Score 5759; DB 9; Length 1140;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1133; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 18 VSTHTSSTLPAPFKKILVANRGEIARAFRAALETGAATVAIYPREDRGSFHRSEAVR 77
Db 1 MSTHTSSTLPAPFKKILVANRGEIARAFRAALETGAATVAIYPREDRGSFHRSEAVR 60

QY 78 IGTEGSPVKAYLIDDEIIIGAARKKADAIYPGYFLSENAQLARECAENGITFTGPTPEV 137
Db 61 IGTEGSPVKAYLIDDEIIIGAARKKADAIYPGYFLSENAQLARECAENGITFTGPTPEV 120

QY 138 LDLTGDKSRANTAAKAGLPVLAESTPSKNIDDIIVKSAEGOTYPIFVKAVAGGGGRMRF 197
Db 121 LDLTGDKSRANTAAKAGLPVLAESTPSKNIDDIIVKSAEGOTYPIFVKAVAGGGGRMRF 180

QY 198 VASPDDELRLKLAESREAEAFGDSVYVERAVINPQIHIEVQIILGDRGTGEVHHLYERDCS 257
Db 181 VASPDDELRLKLAESREAEAFGDSVYVERAVINPQIHIEVQIILGDRGTGEVHHLYERDCS 240

QY 258 LQRRHQKVEIAPAOHLDPDLRDRICADAVKFCRSIGYQAGTGVFEVLVDEKGNHVFIEMN 317
Db 241 LQRRHQKVEIAPAOHLDPDLRDRICADAVKFCRSIGYQAGTGVFEVLVDEKGNHVFIEMN 300

QY 318 PRIQVEHTVTEEVTDVLVKAQMLAAGATLKLGLTQDKIKTHGAALQCRITTEDPNNG 377
Db 301 PRIQVEHTVTEEVTDVLVKAQMLAAGATLKLGLTQDKIKTHGAALQCRITTEDPNNG 360

QY 378 FRPDGTGITAYRSPGAGVRLDGAOLGGEITAHFDSMLVKMTCRGSDFETAVARAQAL 437
Db 361 FRPDGTGITAYRSPGAGVRLDGAOLGGEITAHFDSMLVKMTCRGSDFETAVARAQAL 420

QY 438 AFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEOGRILDYADV 497
Db 421 AFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEOGRILDYADV 480
QY 498 TVNKGVRPKDVAAPIDKLPNTKIDPLPRGSRDLKQLGPAFAFARDLREQDALAVTDTT 557
Db 481 TVNKGVRPKDVAAPIDKLPNTKIDPLPRGSRDLKQLGPAFAFARDLREQDALAVTDTT 540
QY 558 FRAHOSLLATRVRSFALKPAEAAVAKLTPELLSVEAWGATYDVAMRFLFEDPWRLDE 617
Db 541 FRAHOSLLATRVRSFALKPAEAAVAKLTPELLSVEAWGATYDVAMRFLFEDPWRLDE 600
QY 618 LREAMPNVNIQMLLRGNTVGYTPYDPSVCRFAFVKAASGVDFIRFDALNDVDSMRPA 677
Db 601 LREAMPNVNIQMLLRGNTVGYTPYDPSVCRFAFVKAASGVDFIRFDALNDVDSMRPA 660
QY 678 IDAVLETNTAVAEVAMAYSGLSDPNKLYTLDYILKMAEEIVKSGAHILAIKDMAGLLR 737
Db 661 IDAVLETNTAVAEVAMAYSGLSDPNKLYTLDYILKMAEEIVKSGAHILAIKDMAGLLR 720
QY 738 PAAVTKLVTALRREFDLPVHVHTHTAGGOLATYFAAAQAGADAVDGCASAPLSTTQPS 797
Db 721 PAAVTKLVTALRREFDLPVHVHTHTAGGOLATYFAAAQAGADAVDGCASAPLSTTQPS 780
QY 798 LSAIAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL 857
Db 781 LSAIAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL 840
QY 858 SNLRAQATLGLADREFELIEDNTAAVNMGLRPTKVTTPSSKVGDJLALHLVGAGVDPADF 917
Db 841 SNLRAQATLGLADREFELIEDNTAAVNMGLRPTKVTTPSSKVGDJLALHLVGAGVDPADF 900
QY 918 AADPKYDIPDSVIAFLRGELGNPPGCGWPEPLTRALEGRSEKAPLTVPEEEOAHLDA 977
Db 901 AADPKYDIPDSVIAFLRGELGNPPGCGWPEPLTRALEGRSEKAPLTVPEEEOAHLDA 960
QY 978 DSKERRNSLNRLLFPKPTPEEFLEHRRFRFGNTSALDDREFYGLVEGRETLIRLPDVRT 1037
Db 961 DSKERRNSLNRLLFPKPTPEEFLEHRRFRFGNTSALDDREFYGLVEGRETLIRLPDVRT 1020
QY 1038 LVLRLDAISPPDDKGRNVANVNGQIRPMVRDRDSVESVTATAEKADSSNKGHVAAPFA 1097
Db 1021 LVLRLDAISPPDDKGRNVANVNGQIRPMVRDRDSVESVTATAEKADSSNKGHVAAPFA 1080
QY 1098 GVVTVTVAEGDEVKAGDAVAIIEMKMEATITASVDGKIERYVVPAAATKVEGGDLVWVS 1157
Db 1081 GVVTVTVAEGDEVKAGDAVAIIEMKMEATITASVDGKIERYVVPAAATKVEGGDLVWVS 1140

RESULT 6

US-09-815-242-5468
; Sequence 5468, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ ID NO 5468
; LENGTH: 1147
; TYPE: PRF
; ORGANISM: Staphylococcus aureus
US-09-815-242-5468

Query Match 42.4%; Score 2489; DB 10; Length 1147;
Best Local Similarity 45.2%; Pred. No. 3.3e-160;
Matches 527; Conservative 193; Mismatches 383; Indels 62; Gaps 14;

QY 30 KKLIVANRGEIARAFRAALETGAATVAIYPREDGRSFHRSFASAEAVRIGTEGSPVKAYL 89
Db 4 KLLLVANRGEIARIFRAAAELDLSTVAIYSNEDKSLHRYKADESILVSGDLGPAESYL 63
QY 90 DIDEIIAAGKVKADAIYPGYFLSENAQLARECAENGITFIGPTPEVLDLTGDKSAVT 149
Db 64 NIERIIDVAQANVDALHPGCGFLSENEQFARRCAESGIFGPHLEHLMFGDKVKART 123
QY 150 AAKGAGLPLV-ABSTPSKNIDDIIVKSAEGOTYPIFYKAVAGGGGRMRFVSSPDRLKLA 208
Db 124 TAIKADLPVPGTDGPIKSYELAKEFAEAGFPLMIKATSGGGKGMRIVREESELEDAF 183
QY 209 TEASREAAFGDGVVYVRAVINPOHIEVQILGDRPEVHVLYERDCSLORRHKVVEI 268
Db 184 HRKSEAEKSGFSEVYERYIDNPKHIEVOVIGDEHGNIVHLPERDCSVORRHKVVEV 243
QY 269 APAQHLDPPELRDICADAVKFCRSIGYGAGTVEFLV--DEKGNHVFIEHNPRIQVEHTV 326
Db 244 APSVGLSPTLRQICDAAIOLMENIKYVNAAGTVEFLVSGDE--FFIEVNPRIQVEHTI 300
QY 327 TEEVTEVDLVKQMRLAAGATL--KELGLTQDK--IKTHGAALQCRITTEDPNNGRFDTG 383
Db 301 TEMVTGIDIVKTQITLVAAGADLFGEENMPQOKDITTLGVAIQCRITTEDPNDMPDTG 360
QY 384 TIATYRSPGAGVRLD--GAAQLGGEITAHFDSMLVKTKCRGSDPETAVARQALAEFTV 442
Db 361 TIATYRSGGFGVRLDAGDGFQGAETISPYDSSLVLLKLTTHAISPKQAEKMRSLREMRI 420
QY 443 SGVATNIGFLRALIREDEFTSKRIATGFIGDHPHLLQAPPADDEOGRILDYADVTKNP 502
Db 421 RGVKTNPFLINVMKNKFTSGDYTKFIEETPELFDIQSLDRGCTLEYIGNVTIN-- 478
QY 503 HGVKPKDVAAPIDKLPNIKDLPLP-----RGSRLRLKQLGPAAPAFARD 544
Db 479 -----GFPNVEKRPKPDYELASITPVSSSKIASFGTKQLLDEVGKVAEW 525
QY 545 LREQDALAVTDTTFRDAHQSLLATRVRSFALKPAEAAVAKLTPELLSVEAWGATYDVAM 604
Db 526 VKKQDDVLLDTTFRDAHQSLLATRVRTKDMINIASATADVFKDGFSLMMGGGATFDVAY 585
QY 605 RFLFEDPWRLDELREAMPNVNIQMLLRGNTVGYTPYDPSVCRFAFVKAASGVDFIR 664
Db 586 NFLKENPWERLERLRKAIPNVLFQMLLRASNNAVGYKNYPDNVIHFFVQESAKAIDVRI 645
QY 665 FDALNDVDSMRPAIDAVLETNTAVAEVAMAYSGLSDP--NEKLYTLDYILKMAEEIVKSG 723
Db 646 FDSLNVWDQMKVANEAVQEAQ--KISEGTCYTGDLNLPERSNIYTYLEYVYVKLELEREG 704
QY 724 AHILAIKDMAGLLRPAAVTALRREFDLPVHVHTHTAGGOLATYFAAAQAGADAV 783
Db 705 FHILAIKDMAGLLRPAAYELIGELKAAVDLPIHLHTHTDSGNLLTYKQIDAAGVDIID 764


```
Db 1058 SAVQVQKRAEPTNKEIGATMSGSLVQLVLRGDKVKEGQPLITEAMKMETTIEARFAG 1117
QY 1135 KIERVVVPAATKVEGGDLVVV 1156
Db 1118 TVDHIYVEEGEALSSGDLLEEV 1139

RESULT 8
US-09-815-242-12361
; Sequence 12361, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELTRA-011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12361
; LENGTH: 1073
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12361

Query Match 39.2%; Score 2300; DB 10; Length 1073;
Best Local Similarity 44.8%; Pred. No. 1.9e-147;
Matches 489; Conservative 180; Mismatches 360; Indels 62; Gaps 14;

QY 104 DAIPGCGFLSENAQLARCAENGITFPGTPEVLDLTGDKSRVTAANKAGLPVL-AES 162
Db 2 DAHPGCGFLSENEQFARCAEGKTFPGHLEHDMFGDKVKARITAIKADLPVPGTD 61
QY 163 TPSKNIDDIIVKSAEGTPIFYKAVAGGGRGMRFVSSDPDELKRLKLAATESRAEAAGDG 222
Db 62 GPIKSYELAKEAEAGFLPMIKATSGGCGKGMRIVRREESELEDAFHRAKSAEKSFGNS 121
QY 223 SVYVERAVINPOHIEVQILGDRGTGEVVLHYLRDCSLQRRHQKVEITAPQAHLDPELRDRI 282
Db 122 EVVIERYIDNPKHIEVQVIGDEHGNVHLFRDCSVQRHQKVEVAPSVGLSPTLRQRI 181
QY 283 CADAVKFCRISGYOGAGTVEFLV--DEKGNHVFIEMNPRIQVHTVTEVTEVDLVKAQM 340
Db 182 CDAATQLMENIKYVAGTVEFLVSGDE---FFFIENVRVQVHTITENMTGIDIVKTOI 238
QY 341 RLAAAGATL--KELGLTQDK--IKTHGAALQCRITTTEDPNNGFRPDTGCTITAYRSPGAGVR 397
Db 239 LVAAGADLGEELNPNQKQDIITGLVAGIQCRIITTEDPLNDFPDTGTTIAYRSSGGFGRV 298
QY 398 LD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQALAEFTVSGVATNIGFLRALL 456
```

```
Db 299 LDAGDFOGAELISPYDLSLVKLSTHAISFKQAEERKVRSLREMRIRGVKTNIPFLINVM 358
QY 457 REEDFTSKRIATCFICGHPHLLQAPPADDEOGRILIDYLDADTVNKGPHGVPRKDVAPIDK 516
Db 359 KKKKFTSGDYTTTKFIEETPELFDIQPSLDGRGKTLEYIGNVTIN-----G 403
QY 517 LPNIKDLPLP-----RGSRDRLKQLGPAAPFARDLREODALAVTDITTF 558
Db 404 FNVEKRPKPDYELASIPTVSSSKIASFSGTQKQLLDEVGPKGVAEWVKKQDDVLLTDITTF 463
QY 559 RDAHQSLLATRVRSFALKPAEAVAKLTPELLSVEAWGATYDVAMRFIFEDPMDRLDEL 618
Db 464 RDAHQSLLATRVRTKDMINIASKTADVFVKDGFSLMMGGATFDVAYNFLKENPWERLERL 523
QY 619 REAMPNVNLOMLLRGRNTVGYTPYDPSVCRAFAVKEAASSGVDFIRFDALNDVSOHRPAI 678
Db 524 RKAIPNVLFQMLLRASNAVGYKNYPDNVTHKFVQESAKAGIDVFRFDSLNWVDQMKVAN 583
QY 679 DAVLETNTAVAEVAMAYSGDLSDP-NEKLYTLDYILKMAEIEVKSGAHLAIKDMAGLLR 737
Db 584 EAVQEAG-KISEGTICYTGDILNPERSNIYLEYVVKLAKELEREGFHILAIKDMAGLLK 642
QY 738 PAAVTKLVTALRREFDLPVHVHTHDAGGQLATYFAAQAQADAVDGAASAPLSGTTSQPS 797
Db 643 PRAAVELIGELKAAVDLPILHLTHDTSNGNLLTYKQADAGVDIIDTAVASMSGLTSQPS 702
QY 798 LSAIVAFAHTRDRDTCGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTPGRVYRHEIPGGQL 857
Db 703 ANSLYALNGFPRHLRTDIEGMESLSHYWSTVNTYYSDFESDIKSPNTEIYQHMPGGQY 762
QY 858 SNLRQAATAGLADRPDLIEDNYAAVNEMLGRPTKVTTPSSKVVDGLALHLVAGVDPADF 917
Db 763 SNLSQAQSLGLGERDEVDKMYRRVNFLEGDIVKVTTPSSKVVDGMALYVQNDLDEQSV 822
QY 918 AADPKQYDIPDSVIAFLRGLGNPPGGWPEPLRTRALEGRSEKAGPLTEVPEEQAHLDA 977
Db 823 ITDGYKLDFFESVVSFKGEIGOPVNGKNDLQAVILKQE-----ALTARPGEYLEPVD 878
QY 978 DDSKE-----RRNSLRNLLFPKTEFELEHRRRFGNTSALDDREFFGLVVEGRE 1026
Db 879 EKVERLEEEQOGPVTEQDIISYVLKPYVEQYIOTRNQYGNLSLLEDTTFFFGMRNGET 938
QY 1027 TLIRLPDVRTPLLLRLDAISEDDKGMNVVAVNGVQIRPMRVDRSRVESVTATAEKADS 1086
Db 939 VEIEI-DKGRLLIKLETISEPDENGNTIYYAMNQOARRIYIKDENVHTNANVKPAOK 997
QY 1087 SNKGHVAAFPAGVVT-VTVAEQDEVKAGDAVAIIEMKMEATITASVDGKIERYVVPAAAT 1145
Db 998 SNPSHIGAQMPSGVTSEKVSGETVKANQPLITEAMKMETTITQAPFDGVKQVTVNNGD 1057
QY 1145 KVEGGDLVVV 1156
Db 1058 TIATGDLIIIEI 1068

RESULT 9
US-09-767-479-6
; Sequence 6, Application US/09767479
; Patent No. US20010036654A1
; GENERAL INFORMATION:
; APPLICANT: Haselkorn, Robert
; Gornicki, Piotr
; TITLE OF INVENTION: ACETYL-CoA CARBOXYLASE COMPOSITIONS AND
; METHODS OF USE
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
```

```

;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 22-Jan-2001
; CLASSIFICATION: Unknown
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/468,793
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US SN 07/956,700
; FILING DATE: 02-OCT-1992
; APPLICATION NUMBER: PCT/US93/09340
; FILING DATE: 30-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: ARCD:152/KIT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 447 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-767-479-6

```

```

Query Match 17.6% Score 1032.5; DB 10; Length 447;
Best Local Similarity 48.2%; Pred. No. 3.6e-62;
Matches 216; Conservative 71; Mismatches 148; Indels 13; Gaps 4;

QY 29 FKKILVANRGEIATVAFRAALETGAATVAIYPREDRGSHFSEAVRIGTGPVKAY 88
Db 3 FDKILLANRGEIATVAFRAALETGAATVAIYPREDRGSHFSEAVRIGTGPVKAY 88
QY 89 LDIDEITGAAGKVKADAIYPGYGLSENALARECAENGITFTGPTPEVLDLTGDKSRV 148
Db 62 LNIPLNIAALTRNASAIHPGYGLSENALARECAENGITFTGPTPEVLDLTGDKSRV 148
QY 149 TAAGKAGLPVLAES-----TPSKNIDDIVKSAEGQTYPIFVKAVAGGGGRMRFVSSPDE 203
Db 122 ETMQKAGVTPVPGSEGLVETEGLE---LAKDIGVPMIKATAGGGGRMRLVRSPE 177
QY 204 LRKLATEASREAAAFDGSVYVERAVINPOHIEVOILGDRGTGEVHLRYERDCSLQRRHQ 263
Db 178 FVKFLAAQAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 237
QY 264 KVEITAPQAHLDPDLRDRICADAVKFCRSIGYOGAGTVEFLVDEKGNHVFIEKNPRLQVE 323
Db 238 KLEEAPSPALDSDLRKMGQAQVAFNYTGAGTIEFLDRSGQFYEMNTRIQVE 297
QY 324 HTVTTEVTEVDLVKAQMLAAGATLKGELTQDKIKTHGAALOCRTTDPNNGFRPDTC 383
Db 298 HPVTMTGVDLIVETRIQAGRLR---ITQDQVLRGHAIECRINAEDPDHDFRAPG 354
QY 384 TITAYRPGGAGVRLDGAALGGBITAHFDSMLVKMFCRSDRETAVARAQLAEFTVS 443
Db 355 RISGLPPGGVGRIDSHVYTDVQIIPYDLSLIGKLVWGPDRATAINRMKRLRECAIT 414
QY 444 GVATNIGFLRALLREEDFTSKRIATGFI 471
Db 415 GLPTTIGFHORIMENPQFLQGNVSTSFV 442

```

RESULT 10
US-09-767-479-8

```

; Sequence 8, Application US/09767479
; Patent No. US20010036654A1
; GENERAL INFORMATION:
; APPLICANT: Haselkorn, Robert
; Gornicki, Piotr
; TITLE OF INVENTION: ACETYL-CoA CARBOXYLASE COMPOSITIONS AND
; METHODS OF USE
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P. O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09767,479
; FILING DATE: 22-Jan-2001
; CLASSIFICATION: Unknown
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/468,793
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US SN 07/956,700
; FILING DATE: 02-OCT-1992
; APPLICATION NUMBER: PCT/US93/09340
; FILING DATE: 30-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: ARCD:152/KIT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 453 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-767-479-8

```

```

Query Match 17.1% Score 1005.5; DB 10; Length 453;
Best Local Similarity 47.1%; Pred. No. 2.5e-60;
Matches 209; Conservative 72; Mismatches 158; Indels 5; Gaps 3;

QY 29 FKKILVANRGEIATVAFRAALETGAATVAIYPREDRGSHFSEAVRIGTGPVKAY 88
Db 3 FKKILLANRGEIATVAFRAALETGAATVAIYPREDRGSHFSEAVRIGTGPVKAY 88
QY 89 LDIDEITGAAGKVKADAIYPGYGLSENALARECAENGITFTGPTPEVLDLTGDKSRV 148
Db 62 LNIPLNIAALTRNASAIHPGYGLSENALARECAENGITFTGPTPEVLDLTGDKSRV 148
QY 149 TAAGKAGLPVLAESTP-SKNIDDIVKSAEGQTYPIFVKAVAGGGGRMRFVSSPDELRL 207
Db 122 ETMQKAGVTPVPGSEGLTVDVSAKAAEIGYPMIKATAGGGGRMRLVREPDLRL 181
QY 208 ATEASREAAAFDGSVYVERAVINPOHIEVOILGDRGTGEVHLRYERDCSLQRRHQVVE 267
Db 182 FLAAGCEAAAFGNPGLYLEKFDPRHVEFQILADAGYVNVHLGERDCSIORRHOKLE 241
QY 268 IAPAOHLDELDRICADAVKFCRSIGYOGAGTVEFLVDEKGNHVFIEKNPRLQVEHPT 327
Db 242 EAPSPALSDLRKMGQAQVAFNYTGAGTIEFLVDEKGNHVFIEKNPRLQVEHPT 301

```

```

QY 328 EVTEVDLVKAQMRILAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNNGFRPDGTGTTA 387
Db 302 EMITGLDLIAEQIRIAQGAELR---FROADIQLRHAEICRINAEDPEYFNFRNPGRIIG 358
QY 388 YRSPGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDFETAVARAQALAEFTVSGVAT 447
Db 359 YLPPGPGVGVDSHVYTDYEIPYDLSLIGKLVGMATREEAIAARMORALRECAITGLPT 418
QY 448 NIGCFRALLREEDFTSKRIATGFI 471
Db 419 TLSFHQLMQLQMPFELRGELYTNFV 442

RESULT 11
US-09-815-242-5215
; Sequence 5215, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5215
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-5215

Query Match 16.7%; Score 979.5; DB 10; Length 471;
Best Local Similarity 45.9%; Pred. No. 1.5e-58;
Matches 209; Conservative 74; Mismatches 157; Indels 15; Gaps 6;

QY 30 KKILVANRGEIAVRAFRALLETGAATVAIYPREDRGSFHRSEAFSEAVRIGTEGSPVKAYL 89
Db 3 KKILVANRGEIAVRIRACAEIGRSVAVYSEADRHVKKRADEAHSIGAD--PLAGYL 60
QY 90 DIDEILGAKKVKAADATPYGYSFLENALARECAENGITFTGPTPEVLDTGDKSRAYT 149
Db 61 NFRALVNLAVERSCDALHPGYGYSFLENALARECAENGITFTGPTPEVLDTGDKSRAYT 120
QY 150 AAKKAGLPVLAESTSK--NIDDI---VKSAGQYPIPVKAVAGGGGGRMFRVSSPDEL 204
Db 121 SMIAAGVP---CTPGTEGNVADLAELAREAGRIYGVPMKATSGGGGGRIRRCNSREEL 176
QY 205 RKLATEASREAAAGDGSVYVERAVINPOHTEVQILGDRTEGVHLYPERDCSLORRHOX 264
Db 177 EQAFPRVISEATKAFGSAEVEFKECIVNPKHIEAQILADSFNGTVHLFERDCSIQRNOK 236

```

```

QY 265 VVEIAPAOHLDPDLRDRICADAVKFCRSIGYOGAGTVEFLVDEKGNHVFTEMNPRIQVEH 324
Db 237 LIEIAPSPQLTPEQRAYIGDLAVRAAKAVGENAGTVEFLAD-GEVYEMEMNTRVQVEH 295
QY 325 TWTEEVTEVDLVKAQMRILAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNNGFRPDGTGT 384
Db 296 TITEETITGDVVRQIRIASGL---ELSVKQDDIVHRGYALQPRINAEDPKNNFLPSFGK 352
QY 385 ITAYRSPGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDFETAVARAQALAEFTVSG 444
Db 353 ITRYAPGPGVGTDTAIYGTIPTYDYSKMLKLIWALTWEEALDRGLRDLDMRVQ 412
QY 445 VATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQ 479
Db 413 VKTTAPYQEIILRNPEFRSQGNTSFVESHPELTQ 447

RESULT 12
US-09-815-242-13885
; Sequence 13885, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13885
; LENGTH: 449
; TYPE: PRT
; ORGANISM: Salmonella typhi
US-09-815-242-13885

Query Match 16.3%; Score 956; DB 10; Length 449;
Best Local Similarity 47.6%; Pred. No. 5.5e-57;
Matches 206; Conservative 61; Mismatches 158; Indels 8; Gaps 5;

QY 31 KILVANRGEIAVRAFRALLETGAATVAIYPREDRGSFHRSEAFSEAVRIGTEGSPVKAYILD 90
Db 4 KIVIANRGEIALRILRACKELGIKTVAVHSSADRLKHVLLADETVICIGPAPS-VKSYLN 62
QY 91 IDEITGAKKVKAADATPYGYSFLENALARECAENGITFTGPTPEVLDTGDKSRAYT 150
Db 63 IPAILSAEITGAVAIHPGYGYSFLENANFAEQVERSGFIFIGPKADTIRLMGDKVSAITA 122
QY 151 AKKAGLPVLAESTPSKNIDDIKVSACQT--YPIFKVAVAGGGGGRMFRVSSPDELRLKLA 208
Db 123 MKKAGTYTPVGSDDPLGDDMMNANRAHAKRIGYPVIIKASGGGGGGRMFRVRSDAELAQSI 182

```


Qy	29	FKKILVANRGEIAVRAFAALETGAATVATVYIPREDRGSPHRSPASFAVRIGTEGSPKAY	88
Db	2	FRKILIANRGEIAVRIIRAARELGIATVAVYSTADKEALHTLLADEAVCIG-PGRATESY	60
Qy	89	LDIDEIIGAACKYKADAIYPCYGLFSENQAOLARECAENGITFGPTPEVLDLTGDKSRAY	148
Db	61	LNINAVLSAAVLTEAEAIHPGFGFSGNSKAFATMCVEVGIKFIPGSHGVMDMMGDKNINAR	120
Qy	149	TAACKAGLPVLAESTPS-KNIDIDIVASAEQOTYPIPVKAVAGGSGMRFSVSSPDLRLK	207
Db	121	AQMIKACVPVPGSDGGEVHNSEBALIVAEKIGYPMVKASAGSGGKIRKVEKPDDLVSAL	180
Qy	208	ATEASRAEAAFGDGSVYVPRAVINPQHIEVQILGORTGEVHULYPRDCSLORRHQKVVE	267
Db	181	FETASSEKANYONGAMVIERVYIPARHIEVQILGDBEHGVHILGDERDCSLORNNQKVLE	240
Qy	268	IAPAQHILDELRLORICADAVKFCRSIGOGAGTVEFLVDB-KGNHVFIEMNPRIOVEHPV	326
Db	241	ESPSAIGKTLRHEIGAAVRAAEFVGYENAGTIEFLLEASSNFYEMMTRVQVEHPV	300
Qy	327	TEEVTEVDVLKQMRNLAAGATLKEGLGTQDKIKTHGAALOCRTITTEDPNNGFRPDTGTT	386
Db	301	TEFVSGVDVYKEIGICATAAG--QPLSVKQBDIVLRGHAIETCRINAENPAFNEAPSPGKIT	357
Qy	387	-AYRPSGAGVRLDGAQLGGETTAHFDSMLVKWTCRGSDFEFAVARAQALAEFTVSGV	445
Db	358	NLYLPSGCVGLRVDASVYPGYTTPPYDYSMAKIIYVHGNRRFDALMKMQRALYELIEIGV	417
Qy	446	ATNIGFURALLREEDFTSKRIATGFI	471
Db	418	QTNADFOLDLISDRNNIAGDYDTFSFL	443

; GENERAL INFORMATION:
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari L.
 ; APPLICANT: Zyskind, Judith W.

```

, ORGANISM: Streptococcus pneumoniae
, TITLE OF INVENTION: Identification of Essential Genes in
, TITLE OF INVENTION: Prokaryotes
, FILE REFERENCE: ELITRA.011A
, CURRENT APPLICATION NUMBER: US/09/815,242
, CURRENT FILING DATE: 2001-03-21
, PRIOR APPLICATION NUMBER: 60/131,078
, PRIOR FILING DATE: 2000-03-21
, PRIOR APPLICATION NUMBER: 60/206,848
, PRIOR FILING DATE: 2000-05-23
, PRIOR APPLICATION NUMBER: 60/207,727
, PRIOR FILING DATE: 2000-05-26
, PRIOR APPLICATION NUMBER: 60/242,578
, PRIOR FILING DATE: 2000-10-23
, PRIOR APPLICATION NUMBER: 60/253,625
, PRIOR FILING DATE: 2000-11-27
, PRIOR APPLICATION NUMBER: 60/257,931
, PRIOR FILING DATE: 2000-12-22
, PRIOR APPLICATION NUMBER: 60/269,308
, PRIOR FILING DATE: 2001-02-16
, NUMBER OF SEQ ID NOS: 14110
, SOFTWARE: FASTSEQ for Windows Version 4.0
, SEQ ID NO 13617
, LENGTH: 455
, TYPE: PRT
, ORGANISM: Streptococcus pneumoniae
US-09-815-242-13617

```

Query Match 16.0%; Score 936.5; DB 10; Length 455;
Best Local Similarity 45.7%; Pred. NO. 1.2e-55;
Matches 204; Conservative 76; Mismatches 159; Indels 7; Gaps 5;



Qy 209 TEASREAAFGDGVVVERAVINPOHIEVOILGDRTEGVVHLYERDCSLQRRHQKVEI 268
 Db 182 EASREASAFGDPVTVLEQAVINPRHIEVOILADNLGDLVHLYERDCSVORRHQKVEI 241
 Qy 269 APAOHLDELDRKADAVKFCRISGOGATVEFLVDEKGNHVFIEIMNPRIOVEHVTVE 328
 Db 242 APAHDLAELRYKCVDAVAFARIHIGYSCAGTVEFLDERGEYVFIEIMNPRIOVEHVTVE 301
 Qy 329 ETVEDVLKQAMRLAAGATKELGLTQDKTKTHGAALOCRTITDPNNGRPDGTITAY 398
 Db 302 EITVDLVASOLRTAAGTELQGLREDIAPHGAALOCRTITDPNNGRPDGTITAY 361
 Qy 399 RSPGAGVRLDGAALGCEITAHFDSMLVMTKRGSDFTAVARAQALAEFTVSGVATN 448
 Db 362 RTAGAGVRLDGSNLAAGTSPYFDSMLVMTKRGSDFTAVARAQALAEFTVSGVATN 421
 Qy 449 IGLRALREEDFTSKRIATGFIGDHPHLLQAPADDEQGRILDYLDVTVNPKPHGVRPK 508
 Db 422 IPFLQAVLDLDPFRAGRVVTSFIDERPQLLTARASADRGTKILNFLADVTVNPNYGRPS 481
 Qy 509 DVAAPIDKLPNIKDLPL----PRGSRDLKOLGPAFAFARLDREODALAVDTTFRDAHQS 564
 Db 482 TI-YPDOKLP---DLDLRAAPPAGSKORLVKLGPEGFARWLRESAAVGVDTTFRDAHQS 537
 Qy 565 LLATRVRSFALKPAEAAVAKLTPELLSVYEWAGGATYDVAMRFLFEDPWRLDELREAMPN 624
 Db 538 LLATRVRTSGLSRVAPYLARTMPOLLSVECWGGATYVALRFLKEDPWRLDELREAMPN 597
 Qy 625 VNIQMLLRGNVTGTPYVPSVCRFAVKEAASGVDFRIFDALNDVYSQMRPAIDAVLET 684
 Db 598 ICLQMLLRGNVTGTPYVPEIVTSFAVQEAATATGIDIFRIFDALNNTESMRPAIDAVRET 657
 Qy 685 NTAVAEVAMVSGDLSDPNEKLYTLDYLLKMAEIVKSGAHILAIDKMAGLRPAATVKL 744
 Db 658 GSATAEVAMCYTGDLTDPGEOLYTLDYLLKLAEOIVDAGARVLAIDKMAGLRPAATVKL 717
 Qy 745 VTALRRDLPVHVTHTDGTAGOLATFAAAQACADAVDASAPLSGTTSPOLSALVAA 804
 Db 718 VSALRSRDLPLVHLTHDTPGQSLASVAAWHAGADAVDGAAPLAGTTSOPALSSVAA 777
 Qy 805 FAHTRDTPSLEAVSDLEPWEAVRGLYLPFESGTPGTPGRVYRHEIPGQSLNLRQA 864
 Db 778 AAHTFYDGLSLSAVCALEPWEALRYVAFESGLPGTPGRVYRHEIPGQSLNLRQA 837
 Qy 865 TALGLADRFELIDYNAVNEMLRGPTKVTSPSSKVVGDLALHLVAGVDPADFAADPKY 924
 Db 838 TALGLDGRFEEIEAYAGADRVGLRVKVTPTSKVVGDLALALVAGVSADEFAADPARF 897
 Qy 925 DIPDSVIAFLRGELGNPPGWPPELRLTRALEGRSEGKAPLTPVEPEEQAHLDADDSKERR 984
 Db 898 GTPESVLGFLRGELGDPGWPPELRLTAALAGGAAR-PTAQLAADDEIALSSVGAK-RQ 955
 Qy 985 NSLRLPLPKTEFFLEHRRFRGNTSALDDREFYGLVEGRETLIRLPDVRTPLLRDLA 1044
 Db 956 ATLRLPLPPTKEFNEHREAYGDTSQLSANQFFYGLRGEHVRKL-ERGVELLIGLEA 1014
 Qy 1045 ISEPDKGMNVVANVCOIRPMVRDRSPVESVTATAEKADSSNKHVAAPFAGVVTVTY 1104
 Db 1015 ISEPDGCMRVIMLNGQLRPLVLRDSIASAVPAEAKADGNPGHIAAPFAGVVTIVG 1074
 Qy 1105 AEGDEVKAGDAVATIEAMKBEATTASVDGKIERYVPAATKVEGGDLIVVVS 1157
 Db 1075 CVGERVAGQTIATIEAMKNEAPITAPVAGTVVERVAVSDTAQVEGGDLIVVVS 1127

RESULT 2
 A47255
 pyruvate carboxylase (EC 6.4.1.1) precursor [similarity] - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 01-Feb-2002
 C:Accession: A47255
 R:Zhang, J.; Xia, W.L.; Brew, K.; Ahmad, F.
 Proc. Natl. Acad. Sci. U.S.A. 90, 1766-1770, 1993

A:Title: Adipose pyruvate carboxylase: amino acid sequence and domain structure deduc
 A:Reference number: A47255; MUID:93189578; PMID:8446588
 A:Accession: A47255
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1178 <ZNA>
 A:Experimental source: GB:L09192; MID:g293743; PIDN:AAA39737.1; PID:g293744
 A:Cross-references: GB:L09192; MID:g293743; PIDN:AAA39737.1; PID:g293744
 A:Note: sequence extracted from NCBI backbone (NCBIN:126874, NCBIPI:126875)
 C:Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-bindin
 C:Keywords: biotin binding; ligase; mitochondrion
 F:1-20/Domain: transit peptide (mitochondrion) #status predicted <TRP>
 F:21-1178/Product: pyruvate carboxylase #status predicted <MAT>
 F:39-494/Domain: biotin carboxylase homology <BCH>
 F:1105-1178/Domain: lipoyl/biotin-binding homology <LPB>
 F:1144/Binding site: biotin (Lys) (covalent) #status predicted

Query Match 43.5%; Score 2550.5; DB 1; Length 1178;
 Best Local Similarity 46.1%; Pred. No. 2.8e-132;
 Matches 543; Conservative 192; Mismatches 407; Indels 35; Gaps 13;

Qy 7 GGLLLKGIITLVSTHTSSTLPA-----PKKILVANRGETAVRAFRALLETGAA 54
 Db 9 GGLRLGV-----RRSSAPVAPSNVRLEYPKIKKVMVANRGEIATVFRACTELGR 62
 Qy 55 TVAIYIPREDRGSFHRSFASAEVAVRIGTEGSPVKAYLIDIEIIIGAAKKVKADAIYPGYGFLS 114
 Db 63 TVAVYSEQDTGQMHKQKADAEVILGRGLAPVQAYLHPDIIVKAKENGVDVHPGYGFLS 122
 Qy 115 ENAQLARECAENGTIFGPTPEVLDLTGDKSRVATAAKKAGLPVL-AESTPSPKNTDDIVK 173
 Db 123 ERADFAQACODAGVRFIGPSPVVRKMGDKVEARAIAAAGVPVPGTDSPISSLHEAHE 182
 Qy 174 SAGQGTPIFVKAVAGGGRGMRVSSDPDELKRLATEASREAFAGDGVSVVERAVINP 233
 Db 183 FSNTPGFPPIFKAAVGGGGRGMRVHSYELEBENTRAYSEALAAFGNGLFVERFIEP 242
 Qy 234 OHIEVOILGDRTEGVVHLYERDCSLQRRHQKVEITAPAOHLDELDRKADAVKFCRSI 293
 Db 243 RHIEVOILGDRTEGVVHLYERDCSLQRRHQKVEITAPAOHLDELDRKADAVKFCRSI 302
 Qy 294 GYOGAGTVFELVDEKGNHVFIEIMNPRIOVEHVTVEVDLVKQAMRLAAGATKELGL 353
 Db 303 GYENAGTVFELVDEKGNHVFIEIMNPRIOVEHVTVEVDLVKQAMRLAAGATKELGL 362
 Qy 354 TQDKIKTHGAALOCRTITDPNNGRPDGTITAVRSPGGAGVRLDGAAL-QGEITAHF 412
 Db 363 RQENIRINGCATQCVRVTTEDPARSQPDTGRIEVFRSGEGMIRLDNASAFQGVISPHY 422
 Qy 413 DMLVMTKRGSDFTAVARAQALAEFTVSGVATNIGFLRALLREEDFTSKRIATGFIG 472
 Db 423 DSLLVVIAHGDHPTAATKMSRALAEFRVGVKTNIPFLQNVNQQFLAGTVDTQFID 482
 Qy 473 DHPHLLQAPADDEQGRILDYLDVTVNPKPHGVRPKDVA-APID-KLPNIKDLPPLRGSR 530
 Db 483 ENPELFQLRPAONRAOKLLHLYGHVNVNGTTPIPVNVSPSPVDPAPVVPVPGPPAGFR 542
 Qy 531 DRLKOLGPAFAFARLDREODALAVDTTFRDAHOSLQSLATVRSFALKPAEAAVAKLTPELL 590
 Db 543 DILLREGPGFAKAVRNHOGLLMDITFRDAHOSLQSLATVRSFALKPAEAAVAKLTPELL 602
 Qy 591 SVEAWGATYDVAMRFLFEDPWDRDELREAMPNVIQMLLRGNRTVGTTPDPSVCRFA 650
 Db 603 SMENWGATYDVAMRFLYECPPRRRQELRELIPNIPFQMLLRGANAVGTYNPDNVVFKF 662
 Qy 651 VKEAASSGVDFIRIFDALNDVYSQMRPAIDAVLETNTAVAEVAMVSGDLSDPNEKLYTLD 710
 Db 663 CEVAKENGMDVFRVFDLSNLPNMLLGMGAAGSAG-GVVEAAISYTGVDVADPSRTKSYLE 721
 Qy 711 YLLKMAEIVKSGAHILAIKDMAGLLRPAATVKLVTALRREF-DLPVHVTHTDGTAGQOLA 769
 Db 722 YVMGLAEELVRAGTHILCIKDMAGLLKPAACATMLVSSLRDRFPDLPLHITHDTSAGVA 781

QY 770 TYFAAQAQADAVDASAPLSTGTTSPQSLSAIVAFAHTRRTDTGLSLEAVSDLEPYWEAV 829
Db 782 AMLACAQACADVVDVADSMGNTSQPSMGALVACTKGTPLDTEVPLRVEDYSEWEGA 841
QY 830 RGLYLPFESGTGPTCR--VYRHEIPGGQSLNLRQAATAGLADRPDLLELDNYAANVEML 887
Db 842 RGLYAAFDCTATMKSGNSDVYENIPGGQYTNLHFQAHSMGLGSKFKEVKAYVEANQML 901
QY 888 GRPTKVTSPSKVVGDLALHLVAGVDPADFAADPKQYDIPDSVIAFLRCELGNPPGWPE 947
Db 902 GDILKVTSPSKVIGDLAQFWQNGLSRAEAEQAELSFRSVVEFLQYIGIPHGFPPE 961
QY 948 PLRTRAL-----EGRSEKAPLTVPEEEQAHLDA-DDSKERRNSLNRLLPKPTPEEL 1000
Db 962 PFRSKVLKDLPRIEGRPGASLPPLNLKLEKDLIDRHGEVTPEDVLSAAMPDVFAQFK 1021
QY 1001 EHRFRFGNTSALDDREFFVGLVEGRETLLRLPDVTRPLVRLDAISEPDDKGRNVVANV 1060
Db 1022 DFTATFGPLDSNTRLFLQGPKEAEFEVELEGRGT-LHIKALAVSDLNRAGQROVFFEL 1080
QY 1061 NGQIRPMVRDRSVESVTATAEKADSSNKGHVAAPPAG-VVTVTVAEGDEVKAGDAVATI 1119
Db 1081 NQOLRSILVKDQAMKEMHFHFKALKDKVKGQIGAPMPGKVIDIKVAAGDKVAKGQPLCVL 1140
QY 1120 EAMKMEATITASVGGKIERVVVPAATKVEGGDLVVV 1156
Db 1141 SAMKMETVVTSPMEGTIRKVVHTKMTLEGDDLILEI 1177

RESULT 3
JC4391
pyruvate carboxylase (EC 6.4.1.1) precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 20-Jan-1996 #sequence revision 25-Apr-1997 #text_change 11-Jan-2002
C:Accession: S68252; S72393; JC4391; S06440
R:Jitrapakdee, S.; Booker, G.W.; Cassidy, A.I.; Wallace, J.C.
Biochem. J. 316, 631-637, 1996
A:Title: Cloning, sequencing and expression of rat liver pyruvate carboxylase.
A:Reference number: S68252; MUID:96257760; PMID:8687410
A:Accession: S68252
A:Molecule type: mRNA
A:Residues: 1-1178 <JIT1>
A:Cross-references: EMBL:U36585; NID:g1040973; PIDN:AAC52668.1; PID:g1040974
A:Accession: S72393
A:Molecule type: protein
A:Residues: 489-505 <JIT2>
A:Experimental source: liver
R:Lehn, D.A.; Moran, S.M.; MacDonald, M.J.
Gene 165, 331-332, 1995
A:Title: The sequence of the rat pyruvate carboxylase-encoding cDNA.
A:Reference number: JC4391; MUID:96096548; PMID:8522203
A:Accession: JC4391
A:Molecule type: mRNA
A:Residues: 1-221, 'P', 223-865, 'D', 867-976, 'G', 978-1178 <LEH>
A:Cross-references: GB:U32314; NID:g929987; PIDN:AAA96256.1; PID:g929988
A:Experimental source: liver
R:Thampy, K.G.; Huang, W.Y.; Wakil, S.J.
Arch. Biochem. Biophys. 266, 270-276, 1988
A:Title: A rapid purification method for rat liver pyruvate carboxylase and amino acid
A:Reference number: S06440; MUID:89024676; PMID:3178228
A:Accession: S06440
A:Molecule type: protein
A:Residues: 'SG', 23-25, 'PL', 28-29, 'LL', 32-34, 'P', 1134, 'A', 1136-1137, 1139-1152, 'T', 1154-1155
C:Comment: This enzyme is located in the mitochondrial matrix and catalyzes the conversion of pyruvate to oxaloacetate, an intermediate that exit the mitochondrion for consumption in various pathways.
C:Genetics:
A:Gene: pc
A:Genome: nuclear
C:Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binding
C:Keywords: biotin binding; gluconeogenesis; homotetramer; ligase; mitochondrion
F:1-20/Domain: transit peptide (mitochondrion) #status predicted <TNP>
F:21-1178/Product: pyruvate carboxylase #status predicted <MAT>
F:339-494/Domain: biotin carboxylase homology <BCH>

F:1105-1178/Domain: lipoyl/biotin-binding homology <LPB>
F:1144/Binding site: biotin (Lys) (covalent) #status experimental

Query Match 43.5%; Score 2548.5; DB 2; Length 1178;
Best Local Similarity 46.3%; Pred. No. 3.5e-132;
Matches 545; Conservative 188; Mismatches 409; Indels 35; Gaps 13;

QY 7 GGLLLKGIITLVSTHTSSTLPA-----FKKILVANRGEIATVAFRAALTEGAA 54
Db 9 GGLRLGV-----RRSSTAPVSPNVRLEYKPIKVVWVANNRGEIATVFRACTELGR 62
QY 55 TVAIYPREDGRSFHRSFASAEVIRTEGSPVKAYLDIDEIIGAARKVKADAIYPGYGLS 114
Db 63 TVAVYSEQDTGMQRKQKADAEVILGRGLAPQVAYLHPDIKIVAKENGVDVAVHPGYGLS 122
QY 115 ENAQLARECAENGITFIGTPTPEVLDLTGDKSRVTAAKKAGLPVL-AESTPSKNIDIVK 173
Db 123 ERADEAQAQODAGVRFIGSPPEVVRKMGDKVEARATAIAAGVPPVPGTSPNSLHEAHE 182
QY 174 SAEGOTYPIFVKAVAGGGGRGMRVSSPDRLRKLATEASREAAFGDGSVYVERAVINP 233
Db 183 FNTYTGFPFIIFKAAVGGGGRGMRVHSYEELBENYTRAYSEALAAFGALFVEKEIEKP 242
QY 234 OHIEVOILGDRTGEVHLRYERDCSLQRRHQKVVYEIAPAQHLDPDLRDRICADAKVFCRSI 293
Db 243 RHIEVQILGDQYGNILHLRYERDCSIQRRHQKVVYEIAPATHLDPQLRSRLTSDSVKLAKQV 302
QY 294 GYOGAGTVFELVDEKGNHVFIEMNPRIQVHVTVEVDLVKAQMLAAGATLKLGL 353
Db 303 GYENAGTVFELVDEKGNHVFIEVNSRLQVHVTVEITDVLVHAQIHVSEGRSLDGL 362
QY 354 TODKIKTHGAALQCRITTEDPNNGRPDGTITAYSPGAGVRLDGAQAL-GGETAHF 412
Db 363 ROENIRINGCATQCRVTTEDPARSQPDTRTEVFRSGEGMIRLNDASAFQGANVLSPHY 422
QY 413 DSNLVKMTCRGSDFFETAVARAQALAEFTVSGVATNIGFLRLLRDEEDTSKRRIATGFTG 472
Db 423 DSNLVKVIANGKHDPHTAATKMSRALAEFRVGVKTNIPFLQNLVNNQQLAGIVDTQFID 482
QY 473 DHPHLLQAPPADDEOGRILLDYLVNTVKNPHGVPRKQVA-APIDKL-PNIKOLPLPRGSR 530
Db 483 ENPELQLRPAQNRRAQKLLHLGHVNVNGTPTPIPVKVSPPVDPVIVPVVPTGPPAGFR 542
QY 531 DRLKQLGPAFAARLREQDALAVTDTFRDAHQSLIATRVRSFALKPAAEAVALKTPPELL 590
Db 543 DILLREGPEGFARAVNHQGLLLMDTTFDAHQSLIATRVTHDLKKIAPYVAHNPNFL 602
QY 591 SVEANGGATYVAMRFLFDPDWDRLDELREAMPNVNIQMLRGRNTVGTYPYSDSVCRAF 650
Db 603 SIENMGATFDVAMRFLYECPWRLQELRELIPNIPFQMLLRGANAVGVTNYPDNVVKF 662
QY 651 VKEAASSGVDFIERIDALNDVQMRPAIDAVLETNTVAEAVAMAYSGDLSDPNEKLYTLD 710
Db 663 CEVAKENGMDVFRIFDSLNLPLNMLLGMEEAAGSAG-GVVEAAISYTGVDADPSRTYSLE 721
QY 711 YVLKMAEELVKGHAHLAKDMAGLLRPAATVKLTALRREF-DLPVHVHTHTDTAGQLA 769
Db 722 YVGLAEELVRAGTHLCLIKDMAGLLKPAACITMLVSSLRDRFPDPLHHTHTDTSSGVA 781
QY 770 TYFAAQAQADAVDGCASAPLSGTTSPQSLSAIVAFAHTRRTDTGLSLEAVSDLEPYWEAV 829
Db 782 AMLACAQACADVVDVADSMGNTSQPSMGALVACTKGTPLDTEVPLRVEDYSEWEGA 841
QY 830 RGLYLPFESGTGPTCR--VYRHEIPGGQSLNLRQAATAGLADRPDLLELDNYAANVEML 887
Db 842 RGLYAAFDCTATMKSGNSDVYENIPGGQYTNLHFQAHSMGLGSKFKEVKAYVEANQML 901
QY 888 GRPTKVTSPSKVVGDLALHLVAGVDPADFAADPKQYDIPDSVIAFLRCELGNPPGWPE 947
Db 902 GDILKVTSPSKVIGDLAQFWQNGLSRAEAEQAELSFRSVVEFLQYIGIPHGFPPE 961
QY 948 PLRTRAL-----EGRSEKAPLTVPEEEQAHLDA-DDSKERRNSLNRLLPKPTPEEL 1000

A:Cross-references: GB:299111; GB:AL009126; NID:G2633699; PIDN:CAB13359.1; PID:G2633857
A:Experimental source: strain 168
C:Genetics:
A:Gene: PYCA
C:Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binding
F:8-465/Domain: biotin carboxylase homology <BCH>
F:1073-1146/Domain: lipoyl/biotin-binding homology <LPB>
F:1112/Binding site: biotin (lys) (covalent) #status predicted

Query Match 43.1%; Score 2530; DB 2; Length 1148;
Best Local Similarity 46.0%; Pred. No. 3.6e-131;
Matches 533; Conservative 188; Mismatches 390; Indels 48; Gaps 13;

QY 28 AFKKILVANRGEIAVRAFRALTAETGAATVAIYPRDRGSGFHRSEAFSEAVRIGTEGSPVKA 87
DB 5 STOKVLVANRGEIAIRFRACTELNRITVANYSKEDSGSYHRYKADEAYLVGEGKPIDA 64

QY 88 YLDIDIEIIAAGKVKADATPYGGLFSENAQLARECAENGIITFGPTPEVLDLTGDKSRA 147
DB 65 YLDIEGIIIDIAKRNKVDATHPGYGLFSENIHFARRCEEIGVIFGPKSEHLDMEFGDKVKA 124

QY 148 VTAACKAGLPVLAEST-PSKNIDDDIVKSAEGQTYPIFVKAVAGGGGRGMRVSSPDELK 206
DB 125 REQAEKAGIPVIGSDGPAETLEAVFQFGQANGYPIIIRKASLGSGGRGMRIVRSESEVKE 184

QY 207 LATEASREAEAFGDSGVYVERAVINPOHIEVOILGDRTEGVVHLYERDCSLQRRHOKYV 266
DB 185 AYERAKSEAKAFNDENVIEKLIENPKHIEVOIGDKGNVHLFERDCSVQRRHOKVI 244

QY 267 ETAPQHLDELDRICADAVKFCRSIGYOGATVBEFLVDEKGNHVFIEEMNPRIQVETRV 326
DB 245 EVAPSVLSPELRDQICEAAVALAKNVNINAGTVBEFLV-ANNEFYFIEVNPVQVETI 303

QY 327 TEEVTEVDLVKAQMLRAAGATL--KELGLTQDK-IKTHGAALOCRTITTEDPNNRFPDTC 383
DB 304 TEMITGVDIVQIILVAQGLSHLSKKNVNIPEQKDIETIGYATQSRVTTEDPQNDPMDTG 363

QY 384 TITAYRSPGAGVRLD-GAAQLGGEITAHFDSMLVMTCTGSGDFETAVARAQRLAEFTV 442
DB 364 KIMAYRSGGGFVRLDTGNSFGAVITPYDLSLVKLTWALTFEQAQAAKKNVNLQEFRI 423

QY 443 SGVATNIGELRALLREEDTSRIATGFTGDHPHLLQAPPADDEQGRILDYADVTVN-- 500
DB 424 RGIKTNIPELVNAKHEKLTGQYDSFIDTTPPELFPKQKDRGKTMLTYIGNVTVNGF 483

QY 501 -----KPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAFAFARDLREQDALAV 553
DB 484 PGIGKKEKPAFDKPLGVKVDVQDP-----ARGTKQILDEKAGBGLANWVKRQKSVLL 536

QY 554 TDTTFDRAHQSLATRVRSFALKPAAEAVAKLTPELLSVEMGATYDVAMRFLFEDPDWD 613
DB 537 TDTTFDRAHQSLATRIHRSHDLKIANPTAALWPELFSMEMGGATFDVAYRFLKEDPWK 596

QY 614 RLDELREAMPNNVNIQMLLRNTGVTPYDPSVCRFAVFEAAASSGVVDIFRFDALNDVSO 673
DB 597 RLEDLRKEVPTLFQMLLRSSNAGVTNYPDNVKEFKVQSAQSGDVFRIEDSLNNVKG 656

QY 674 MRPAIDAVLETNVAEAVAMAYSGDLSDPNEKLYTLTDYVLKMAEEIVKSGAHILAKDMA 733
DB 657 MTLAIDAVRDTG-KVAEAAICYTGDILDKNRTKYDIATYTSMAKEALEAAGAHILGKDMA 715

QY 734 GLLRPAAVTKLVTALRREDFLPVHVHTHTDAGGQATYFAAAQAGADAVDGASAPLSGTT 793
DB 716 GLLKPOAAVELYSALAKETIDIPVHLHTDTSNGIYMYAKAVEAGVDIIDVAVSSNAGLT 775

QY 794 SQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPVWEAVRGVLYLPESCTPGTGRVYRHEIP 853
DB 776 SQPSASGCFHAMEGNDRRPMNVQVGLLSQYVESYKYYSEFSGMKSPHTEIYHEMP 835

QY 854 GGOLSNLRAQATAGLADRFELIEDNAAVNEMLGRPTKVTPTSSKVVGDALHLVAGVD 913
DB 836 GGOYSNLQOQAGVGLGDRWNEVKEMYRRVNDMFGDIVKVTPTSSKVVGDMAVYMNQNNLT 895

QY 914 PADFAADPOKYDIPDSVIAFLRCELGNPPGWPPEPLRTRALEGRSECKAPLTVPEP--- 970
DB 896 EKDVYKESGLDPPDSVVELFKNIGQPHGFFPEKQKLKLGQE-----PIYVRPGELLE 951

QY 971 -----EQAHLDDADDKERNLSNRLFLFPKPTPEEFLEHRRFRFGNTSALDDREFF 1018
DB 952 PVSFEALKQEFKEHNLISD-----QDAVAYALYPKVFTDYVKTTSYGDISVLDTPTF 1007

QY 1019 YGLVEGRETLRLPLDVTRTLLVRLDAISEPDDKGMNVANVNGQIRPMVRDRSVESVT 1078
DB 1008 YGMTGEEIEVEVERGKT-LIVKLISIGEPQDPATRVVYFELANGQPREVVIKDESKSV 1066

QY 1079 ATAEEKADSNKGVAAFPAGVTVTVAE-GDEVKAGDAVAIIIEAMKMEATITASVDGKIE 1137
DB 1067 OERLKADRTNPSHIAASMPGTVIKVLAEAGTKVKNKGDLHLMINAMKMETTVQAPFSGTIK 1126

QY 1138 RVVVPAAATKVEGGDLIVVV 1156
DB 1127 QVHVKNGEPIQTGDLLEI 1145

RESULT 6

JC2460
pyruvate carboxylase (EC 6.4.1.1) precursor - human
N:Alternate names: pyruvate:carbon dioxide ligase (ADP-forming)
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 18-Aug-2000 #text_change 01-Feb-2002
C:Accession: G01933; JC2460; B27883; S01469
R:Walker, M.E.; Jitrapakdee, S.; Val, D.L.; Wallace, J.C.
submitted to the EMBL Data Library, July 1995
A:Reference number: H00708
A:Accession: G01933
A:Status: translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-1178 <WAL>
A:Cross-references: EMBL:U30891; NID:g1101028; PIDN:AAAB2937.1; PID:g1101029
R:MacKay, N.; Rigat, B.; Douglas, C.; Chen, H.S.; Robinson, B.H.
Biochem. Biophys. Res. Commun. 202, 1009-1014, 1994
A:Title: cDNA Cloning of human kidney pyruvate carboxylase.
A:Reference number: JC2460; MUID:94324922; PMID:8048912
A:Accession: JC2460
A:Molecule type: mRNA
A:Residues: 1-224,'wp',227-351,'A',353-384,'PT',387-485,'DV',488-637,'R',639-728,'A',
A:Cross-references: GB:S72370; NID:G632807; PIDN:AAB31500.1; PID:G632808
R:Lambonvah, A.M.; Quan, F.; Gravel, R.A.
Arch. Biochem. Biophys. 254, 631-636, 1987
A:Title: Sequence homology around the biotin-binding site of human propionyl-CoA carb
A:Reference number: A27883; MUID:87212051; PMID:3555348
A:Accession: B27883
A:Molecule type: mRNA
A:Residues: 1083-1178 <LAW>
A:Cross-references: GB:M26122; NID:g189657; PIDN:AAA36423.1; PID:g387003
R:Freytag, S.O.; Collier, K.J.
J. Biol. Chem. 259, 12831-12837, 1984
A:Title: Molecular cloning of a cDNA for human pyruvate carboxylase. Structural rel
A:Reference number: S01469; MUID:85030380; PMID:6548474
A:Accession: S01469
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1135-1178 <FRE>
A:Cross-references: EMBL:K02282; NID:g189655; PIDN:AAA60033.1; PID:g189656
C:Genetics:
A:Gene: GDB:PC
A:Cross-references: GDB:119472; OMIM:266150
A:Map position: 11q11-q13.1
C:Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binc
C:Keywords: biotin binding; gluconeogenesis; ligase; mitochondrion
F:1-20/Domain: transit peptide (mitochondrion) #status predicted <TNP>
F:21-1178/Product: pyruvate carboxylase #status predicted <NAT>
F:39-494/Domain: biotin carboxylase homology <BCH>
F:1105-1178/Domain: lipoyl/biotin-binding homology <LPB>
F:1144/Binding site: biotin (lys) (covalent) #status predicted

Query Match	43.1%	Score	2528.5	DB 1	Length	1178	
Best Local Similarity	45.9%	Pred. No.	4.5e-131				
Matches	540	Conservative	187	Mismatches	415	Indels	35
							13
QY	7	GGLLKGIITLSTVHTSSPLA-----	FKILVANRGEIAVRAFRALLETGAA	54			
DB	9	GGRLGLI-----	RTSTAPAPASPNVRLEYPKIKVMVANRGEIAIRVFRACTELGR	62			
QY	55	TVAIYPREDRGSHRSFASAEAVRICTEGSPVKAYLIDIELIIGAALKVKADAIYGYGFLS	114				
DB	63	TVAIYSEQDTGMHROKADAEVLIIGRLAPQVAYLHPDILIKVAKENNVDVAFHGYGFLS	122				
QY	115	ENNAQLARECAENGIYFIGTPPEVLDTGDKSRVTAAKKAGLPVL-AESPTSKNIDDIVK	173				
DB	123	ERADFAAQODAGVRFIGTPEVPEVRKMGKVEARAIAATAAGVPVPGTDAPITSLHAHE	182				
QY	174	SAGQYTFIVKAVAGGGGRMRFVSSDELKRLATEASREAEAAFGDGSVYVERAVINP	233				
DB	183	FSNTYGFPIFKAAYGGGRMVRVHSVEELEENTRAYSEALAAFGALFVEKFEKP	242				
QY	234	OHIEVOILGDRTGVEVHLYERDCSLORRHQKVEITAPAQHLDPELRDICADAVKFCRSI	293				
DB	243	RHIEVOILGDOYGNILHLYERDCSORRHQKVEITAPAAHLDPOLRLTSDSVKLAKQV	302				
QY	294	GYOAGCTVFLVDEKGNHVFIEEMNPRIQVEHTVEEDVLVKAQMLAAGATLKEGL	353				
DB	303	GYENACTVFLVDRHGKHVFIEVNSRLQVEHTVEITDVLVHAQIHVSEGRSLPDLGL	362				
QY	354	TQDKIRTHGAALQCRITTEDPNNGFRPDGTITATVRSPPGAGVRLDGAAQL-GGEITAHF	412				
DB	363	ROENIRINGCAIQCRVTTEDPARSFQDPTGRIEYFVRSGEGMIRLDNASAFOGAVISPHY	422				
QY	413	DSMLVKWTCRGSDFTAVARAORALAEFTVSGVANNIGFLRALLREEDFTSKRTATGFIG	472				
DB	423	DSLLVAVIAHGRDHPAATKMSRALAEFVRGVKTNIAFLQNLVNLNQOFLAGTVDQTQID	482				
QY	473	DHPHILOAPPADDEQGRILDIYLDADVTKNPKGVRP-KOVAAPIDK-LPNIKDLPLRGSR	530				
DB	483	ENPELQLRPAQNRQAOKLHYLGHWVNGTPTPIPVKASPSPTDPVPVAFVIGPPAGFR	542				
QY	531	DRUKLQGPAAFRDLREQDALAVTDTTPFDAHQSLATRVSRFALKPAEAAVAKLTPELL	590				
DB	543	DILLREGPGFARAVNRHPLGLLMDTTPFDAHQSLATRVKTHDLKLTAPVVAHNFSKLF	602				
QY	591	SVEAWGGATYDVAMRFLFEDPMDRLDELREAMPNVNIQMLLRGNTVGYTYPDPSVCRAF	650				
DB	603	SMENWGAGTFDVAMRFLYPCPWRRLQELRELPNTPFQMLLRGANAVGYTYNPONVYFKF	662				
QY	651	VKEAASGVGDIFRFDALNDVQMRPAIDAVLETNTAVAEAMAYSGDLSDPNEKLYTLD	710				
DB	663	CEVAKENGMDVFRVPDSLNTLPNMLLGMAAGSAG-GVVEAAISYTGVDVADPSRTKYSLQ	721				
QY	711	YLLKMAEETVKSAGAHILAKDMAGLLRPAAVTKLYTLALREF-DLPVHVHTDHTAGGOLA	769				
DB	722	YWGCLAEELVRAGTHILCKDMAGLLKPTACTMLYSSLRDRFPDPLPLHIHTDHTSGAVA	781				
QY	770	TYFAAQAQADAVDASAPLSCTTSQPSLSATVAFAFTRDRTGLSLEAVSDLEPYEAV	829				
DB	782	AMLACAQADVVDAWDSNGMTSQPSMGALVACTRGTPLDTEVPMERVFDYSEYHEGA	841				
QY	830	RGLYLPFESGTPGPTGR--VYRHEIPGGOLSNLRAQATGALGLADRPFLIEDNYAAVNML	887				
DB	842	RGLYAADFCTATMKSGNSDYVNEIPGGYTNLHFQAHSMGLSGFKVKKAYVEANOML	901				
QY	888	GRPTKVTTPSKVYGDIALHLVAGVDPADPADPOKRYDIPDSVIAFLRGELGNPPGGWPE	947				
DB	902	GDLIKVTTPSKVIGDLAQFWONGLSRAEAAQAABELSPFRSVBVFLOGYGVPHGGFPE	961				
QY	948	PLRTRAL-----EGRSEKAPLTVPEPEQOAHLLA-DDSKERRNSLNRLLPKPTPEFL	1000				
DB	962	PFRSKVLKDLPRVEGREGASLPDLLOAEKEKELVDRHGEENTVEDVLNAAMYVDVFAHFK	1021				

Query Match	42.5%;	Score	2491;	DB	2;	Length	1146;																					
Best Local Similarity	45.9%;	Pred.	No. 5e-129;																									
Matches	527;	Conservative	183;	Mismatches	409;	Indels	30;																					
Gaps																												
QY	30	KKILVANRGEI	AVRAFAAL	ETGAAT	VAIY	PREDG	SFHRSE	FASEAV	RGTEG	SPVKAYL	89																	
DB																												
QY	5	KKVLVANRGEIA	IRVRAC	TELKIK	TAIY	SOED	TGSPHRY	KSDRAYL	VGAGK	KPIDAYL	64																	
DB																												
QY	90	DIDIBII	GAARKVK	RADAIY	PGYF	LSENAQ	LARECA	ENGTTF	IGCTP	PEVL	DLTGD	KSRVAT	149															
DB																												
QY	65	DIENII	IAKESG	ADAIHP	CGYF	LSENIE	EFARRC	EQEGIF	IVGPK	SKHLD	MGDKIK	AKKE	124															
DB																												
QY	150	AAKKA	GLPVLA	EST-PSK	NIDDI	TVKSA	EQGTYPI	FIVK	VAGG	GGGRM	FVSP	SDEB	KLKA	208														
DB																												
QY	125	QALLAD	IPITPG	SDGVP	AGVKE	VEFGE	KGNKPYL	MIKAS	LG	GGGRM	FVRESKE	HVEK	SEF	184														
DB																												
QY	209	TEASRE	AAAFG	DGVS	VYVERA	VINPO	HEIO	ILG	DRTG	EVVHL	YERD	CSLQ	RRHOK	VVEI	268													
DB																												
QY	185	ERASSE	AKAAG	FNDE	VYVEK	VNPNR	PHIE	EVQ	ILG	DTHG	NIHV	LH	FERD	CSIQ	RRHOK	VVEV	244											
DB																												
QY	269	APAQH	LDPE	LRD	ICAD	AVK	FCRS	IGY	GAGT	VEFL	VDK	GNHV	FIEM	NPRI	QI	VEHT	TYTE	328										
DB																												
QY	245	APCNA	ITSEL	NRNIC	DAAV	KLM	KNDV	INAG	TVE	FLV	-EGD	NFY	FIEV	NP	RV	QV	EH	TE	303									
DB																												
QY	329	EVT	EDVL	VKAQ	MLA	AGAT	LKEL	GLT	--ODK	IK	THGA	LQ	CRIT	ED	PN	NG	FR	PD	TGTI	385								
DB																												
QY	304	MITG	IDIV	SQLF	IAD	GYAL	HDQ	VAIP	KQED	IH	IGSA	IQSR	IT	ED	PLN	NF	MD	TG	VR	363								
DB																												
QY	386	TAYR	SPG	AG	VR	LD-GAA	Q	LG	GGIT	AH	FS	MLV	KMT	CRG	SD	FET	AVAR	AQ	ALAE	FT	VSG	444						
DB																												
QY	364	DTYR	STG	GG	VR	LDAG	NG	FGT	VVT	PFY	DS	LLV	K	LC	TG	WT	EQ	AT	RK	MRR	N	LIE	FR	ING	423			
DB																												
QY	445	VATN	IG	FL	R	ALL	RE	ED	FT	SK	RIAT	G	IG	DH	PH	LQ	APP	AD	BQ	GR	IL	D	YAD	V	YK	NK	PHG	504
DB																												
QY	424	VKTNI	PELL	N																								

Qy	661	VLEINTIAVAELVAVATLSDSDPPNEKLI	LDITLNMREELIVKSGAHLIALILNMGALTRFAA	7474
Qy	662	VREAG-KVVEATTCTYGDIDDDTRTKYIT	IDYKDMAKELVAQGHILGKONAGLLKQPA	720
Db				
Qy	741	VTKLVTALRREFDLPHVHITHDTAGCOLAT	YFAAQAGADAVDGCASAPLSGTSOPSLSA	800
Qy		:	:	
Db	721	AYRLIGELKDTVDPIHLHITHDTSNGING	IYTAAYASAGVDIVDVASSAMSGATSPSWTG	780
Qy	801	IYAAFAHTRRDTGLSLEAVSDLEPEYWEA	VRGLYLPPESTPGTGTRVRYRHEITPGQSLNL	860
Qy		:	:	
Db	781	LYTGLVNGNRQTNLDAQNSQII	INHYYEDVRHYKYKDFDNALNSPQTEVYI	THEMPGQGYTNL 840
Qy	861	RAQATALGLADRFELTEDNYAAVNEML	GRPTKVTPTSSKVVGGDLAHLVAGVDPDAFAAD	920
Qy		:	:	
Db	841	QQQAIVGLGRDEWEKEMTYVVMQMGDI	IVKVTSSKVVGGDLALFWQONELTEEDVEYK	900
Qy	921	POKYDYDPSYIAFLRGEALNPPGCGWPE	PLRTRALEGRSEKGAPLTEVP-----BEE	971
Qy		:	:	
Db	901	GDTIDFPDSVIEFPMGEIGOPYGGPF	EKLQKLKGRGRT---PLADRP	GALLMEPVNFAEV 956
Qy	972	QAHLDDADSKE--RRNSLNRLPLFPK	PTEEFLEHRRRFGNNTSALDREFFYGLVGR	ETLI 1029
Qy		:	:	
Db	957	KAELEKMGVEPSEKDVISILYKPVFLVD	YQEMISIKYGDVTVLDTPTFFYKGRGLGET	IEV 1016
Qy	1030	RLPDVRTPLLVRLDAISEPDDKGM	RNVNVAWNGSQIRPMRVRDRORSVESV	TATAEKADSSNK 1089

QY 1030 RLPDVRTPLLVRLDALTSEPDDKGNRVVNVANVNGQIRPMRVDRDRSVESVTATAEKADSSNK 10899

Db 617 TEDPWERLSLIREGAPNLLQLRGANGVKNYPDNVVKYFVRQAAAGGVDLFRVDC 676
Qy 668 LNDVSQMRPAIDAVLETNATAVAEMAYSGDLSDPNEKLYTLDYLLKMAEIEVKSAGHIL 727
Db 677 LNWVENMVSMDATAEEN-KICEATICVTGDLNLSARPKYDKLYTTLNLAVELEKAGAHII 735
Qy 728 AIKDMAGLLRPAAVTKLVTLARRERFDLPVHVHTHTDAGGQLATYFAAAQAGADAVDGASA 787
Db 736 AVKDMAGLLKPAKAAKLVFKALREATGLPIHPTHDTSGISAATVLAADAVDAVDAAMD 795
Qy 788 PLSGTTTSPSLSAIIVAAFAHTRRDTGLSLAASVLEPWEAVRGLYLPFESGTPGPTGRV 847
Db 796 AFSGNTSQCLGSIVEALSSESDTGLDTEWIRISFYWEAVRNOYAFESDLKGPASEV 855
Qy 848 YRHETPGGOLNLRQAATLGLADRFELIEDNYAAVNEMLGRPKTKVPSKVVGDIALHL 907
Db 856 YLHEMPPGQFTNLKEQARSGLSEWRHEVAQYADANRMFGDIVKVPSPKVVGDMLMM 915
Qy 908 VGAGVDPADFAADPKYDIPDSVIAFLRGELGNPPGWPPELRTRALEGRSEKAPLTV 967
Db 916 VSQDLTVADVENPREVSFDPVSMKGLDGLQSPGGWPEALQKKAL----KGEKPYTVR 971
Qy 968 PEE--EQAHLDDADSKERNLSNR-----LLFPKTEEFLEHRRRFGNTSALDDR 1015
Db 972 PGSLEDDADLDA-ERKVIETKLERKVDDEFAFASYLMYPKVFTDFALTAEYGPVSVLP 1030
Qy 1016 EFFYGLVEGRETLIRLPDVRTPLVRLDAISEPDDKGMNRYVANVGOIRPMVRDRS-V 1074
Db 1031 AYPFGMEDGEELFADIERGKTLVIVN-QASSGIDDKGMVTVFEINGOPRRIKVPDRAHG 1089
Qy 1075 ESVTATAEKADSSNKGHVAAFPAGVVT-VTVAGDDVSKAGDAVAILIEMKMEATITASVD 1133
Db 1090 ASGSVRRKAEPCGNASHICAMPGVISRVFINOGQEVKAGDVLLSIEAMKMETALHAERD 1149
Qy 1134 GKIERVVVPAATKVEGGDLIV 1154
Db 1150 GKIAEVLKPGQDIDAKDLII 1170

RESULT 12
AE2911
pyruvate carboxylase [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002
C:Accession: AE2911
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, H.;
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan,
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AE2911
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1174 <KUR>
A:Cross-references: GB:AF008688; PIDN:AAL43707.1; PID:g17741236; GSPDB:GN00186
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: pyca
A:Map position: circular chromosome
C:Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binding

Query Match 42.2% Score 2474.5; DB 2; Length 1174;
Best Local Similarity 46.5% Pred. No. 4.2e-128;
Matches 540; Conservative 174; Mismatches 406; Indels 41; Gaps 19;

Qy 25 TLPAFKTKLVNARGEIARVAFRALETKAATVAILYPREDRGSFHRSEAFSEAVRIG----- 79
Db 20 TVLKISKILVANRSEIARVFRANELGIKTVAIWAEDKLSLHRFKADESYQVGRPHL 79
Qy 80 -TEGSPVKAYLIDDEIGAANKVKADAILYPGYFLSENAQLARCAENGLTFIGPTPEVL 138

Db 80 AKDMGIESYLSSEEVIRVAKLSGADAIHPGYGLLSSEFSEFVACNKAAGITIGPTPTDM 139
Qy 139 DLTGDKSRVTAAKKAGLPVLAESTP-SKNIDDIVKSAEQTYPIFVKAVAGGGGRGMR 197
Db 140 RQLGNKVAARNLAISVDVPVVPATNPLPDIAEVEERMAEIGVPVMLKASWGGGGRMA 199
Qy 198 VSSPDELRLKLAESAEAAFGDGSVYVERAVINPOHIEVOILGDRTEVTVHLEDCS 257
Db 200 IRKKEDLREVTEAKREAAFGKDEVLKLVERRARHVESOILGDTGNVVLPERDCS 259
Qy 258 LORRHOKVVEIAPAOHLDELPLDICADAVKFCRSIGYOGAGTVPELVD-ERGNHVFEM 316
Db 260 IORRNQKVVVERAPAPYLSEAQRQELAAYSILKTAATNYICAGTVEYLMADATGKFYIEV 319
Qy 317 NPRIQVHVTEVTEVDLVKQAMRLAAGATL--KELGL-TQDKIKTHGAALOCRTITTD 373
Db 320 NPRIQVHVTEVTEVDLVKQAHILEGAAIGTAEAGVPKQEDIRLNGHALOCRTITTD 379
Qy 374 PNNGRPDGTITAYRSPGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFFETAVAR 432
Db 380 PEHNFIPDYGRITAYRSASGFIRLDGGYSYTGAVITRYIDPLLVKVTAWAPEPDEAIR 439
Qy 433 AQALAEFTVSGVATNIGFLRALLREEDFTSKRIATGFIQDHPHLLQAPPADBOGRILD 492
Db 440 MDRLREFRIRGVATNLTFLEAIIGHDSFRNNTYTRFDISTPELPAQYKRODRATKLLT 499
Qy 493 YLADVTN--KPHG-VRPKDVAA-PIDKLPNIKDLPLPRGSRDRLLKQLGPAFAARDLRE 547
Db 500 YLADVTNNGHPETKGRAPKADKAAPKPI--VPYI-DAPTQDGTQLLDKLGKGFADMMN 556
Qy 548 QDALAYTDTTFRDADHOSLLATRVRSFALPAAEAVALKTPELLISVBAWGATYDVAMREL 607
Db 557 EKRVLVTDTMRDGHOSLLATRVTHDIARVASVYKALPQLLSLSCWGATFDVSMREL 616
Qy 608 FEDPWRDLDELREAMPNVIQMLLRNRTVGYTPYPSVCRAFVKEAASGVDIFRIDA 667
Db 617 TEDPWERLSLIREGAPNLLQLRGANGVKNYPDNVVKYFVRQAAAGGVDLFRVDC 676
Qy 668 LNDVSQMRPAIDAVLETNATAVAEMAYSGDLSDPNEKLYTLDYLLKMAEIEVKSAGHIL 727
Db 677 LNWVENMVSMDATAEEN-KICEATICVTGDLNLSARPKYDKLYTTLNLAVELEKAGAHII 735
Qy 728 AIKDMAGLLRPAAVTKLVTLARRERFDLPVHVHTHTDAGGQLATYFAAAQAGADAVDGASA 787
Db 736 AVKDMAGLLKPAKAAKLVFKALREATGLPIHPTHDTSGISAATVLAADAVDAVDAAMD 795
Qy 788 PLSGTTTSPSLSAIIVAAFAHTRRDTGLSLAASVLEPWEAVRGLYLPFESGTPGPTGRV 847
Db 796 AFSGNTSQCLGSIVEALSSESDTGLDTEWIRISFYWEAVRNOYAFESDLKGPASEV 855
Qy 848 YRHETPGGOLNLRQAATLGLADRFELIEDNYAAVNEMLGRPKTKVTPSKVVGDIALHL 907
Db 856 YLHEMPPGQFTNLKEQARSGLSEWRHEVAQYADANRMFGDIVKVPSPKVVGDMLMM 915
Qy 908 VGAGVDPADFAADPKYDIPDSVIAFLRGELGNPPGWPPELRTRALEGRSEKAPLTV 967
Db 916 VSQDLTVADVENPREVSFDPVSMKGLDGLQSPGGWPEALQKKAL----KGEKPYTVR 971
Qy 968 PEE--EQAHLDDADSKERNLSNR-----LLFPKTEEFLEHRRRFGNTSALDDR 1015
Db 972 PGSLEDDADLDA-ERKVIETKLERKVDDEFAFASYLMYPKVFTDFALTAEYGPVSVLP 1030
Qy 1016 EFFYGLVEGRETLIRLPDVRTPLVRLDAISEPDDKGMNRYVANVGOIRPMVRDRS-V 1074
Db 1031 AYPFGMEDGEELFADIERGKTLVIVN-QASSGIDDKGMVTVFEINGOPRRIKVPDRAHG 1089
Qy 1075 ESVTATAEKADSSNKGHVAAFPAGVVT-VTVAGDDVSKAGDAVAILIEMKMEATITASVD 1133
Db 1090 ASGSVRRKAEPCGNASHICAMPGVISRVFINOGQEVKAGDVLLSIEAMKMETALHAERD 1149
Qy 1134 GKIERVVVPAATKVEGGDLIV 1154
Db 1150 GKIAEVLKPGQDIDAKDLII 1170

Db 1150 GKIAEVLVPGDQIDAKDILLI 1170

RESULT 13

AE3285

pyruvate carboxylase (EC 6.4.1.1) [imported] - Brucella melitensis (strain 16M)

C:Species: Brucella melitensis

C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002

C:Accession: AE3285

R:DelVecchio, V.G.; Kapral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, E.; Mazur, M.; Goltzman, E.; Selkov, E.; Elizer, P.H.; Hagius, S.; O'Callaghan, D.; Letessier, J.; et al. Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis

A:Reference number: AD3252; PMID:11756688

A:Accession: AE3285

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1158 <CUR>

A:Cross-references: GB:AE008917; PIDN:AAL51448.1; PID:g17982157; GSPDB:GN00190

A:Experimental source: strain 16M

C:Genetics:

A:Gene: BME10266

A:Map position: 1

C:Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binding

C:Keywords: ligase

Query Match 42.0%; Score 2465.5; DB 2; Length 1158;

Best Local Similarity 46.5%; Pred. No. 1.3e-127;

Matches 537; Conservative 172; Mismatches 408; Indels 37; Gaps 15;

QY 31 KILVANRGEIARAFRAALETGAATVAIYPRDRSGFHRSFASEAVRIG-----TEGSP 84

Db 10 KILVANRSEIARVFRANAEGLKTVTWAEDKLSLHRFKADESYQVGRGPHLDRLGP 69

QY 85 VRAYLIDIDEIGNAKKVKAADAYPGFSLSENNALRECAENGIPTGTPVEVLDTGDK 144

Db 70 IESYLIDEIRKVLGSADATHPGYLLSESPFEFAEACAGIVFPGKPTMRLNGK 129

QY 145 SRVATAAKKAGLPVLAESTP-SKNIDDIVKSAEGQYPIFVRKAVAGGGGGRGMRFEVSSPB 203

Db 130 VAARNLAIIEIGVVPVATDPLDDMDVEVKLAAQIGYPLMLKASGGGGRGMRIRAEAD 189

QY 204 LRKLATEASREAAFGDGSVYVERAVINPOHIEVOILGDRGTGEVYVHLVERDCSLQRHQ 263

Db 190 LAREVMEAKREAAFGDKDEVLEKLVARHVEVOILGDTYGNVHLFERDCSIOQRNQ 249

QY 264 KVEETAPAOHLPELDRICADAVKFCRSIGYQAGTVEFLVD-EKGNHVFTEMNRIQV 322

Db 250 KVERAPAPYLNDAAQRELADYGLKTAHATDYIGAGTVEFLMDADYTKFYFIEVNPRIQV 309

QY 323 EHTVTEEVTEVDLVKAQMLAAGATL--RELGL-TQDKIKTHGAALQCRITTEDPNNGFR 379

Db 310 EHTVTEEVTEVDLVKAQMLAAGATL--RELGL-TQDKIKTHGAALQCRITTEDPNNGFI 369

QY 380 PDTGTTATVRSGGAGVRLD-GAQLGGGEITAHFDSMLVKMTCRGSDFETAVARAQALA 438

Db 370 PDYGRIOAYRSAAFGIRLDGGTAYSGAFITRYDPLLVKVTASGATPLEATHRMDRALR 429

QY 439 EFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEOGRILDYADV 498

Db 430 EFRIRGVAATNLFLFAIINHPLKSNIDYTRFDITDTPPELFQMKQDRATKLLTYIADYT 489

QY 499 VN-----KPHGVPRKDDVAAPIDKLPNIKPLPRGSRDLKOLGPAFAFARLDREODALAV 553

Db 490 VNGHPETKGRKAPDAAP--RVWFGDKLVADGTQKLLDQLGPKKFAEWRNKRALI 547

QY 554 TOTTFDAHQSLLATVRFSALKPAEAAVKLTPPELLSVAMGGATYDVAMRLEFEDPDW 613

Db 548 TOTTFDRGHQSLLATVRTYDIARIANAYAAQALPNLFLSECVGGATFDVSMRFLTPDWE 607

QY 614 RLDELREAMPNVNIQMLRGRVTGYTPYDPSVCRAFVVEAASSGVDFIIFRFDALNDVSO 673

Db 608 RLALVREGAPNLLQLLRGANGVGYKSYPDNVVVKYFVREAAARAGIDLFRVFDLSNVEN 667

QY 674 MRPAIDAVLETNTVAEAMAYSGDLSDPNEKLYTLDYYLKMABEIVKSAHILAIKDMA 733

Db 668 MRVSMADVLEEN-KICEAAICTYGDILNPDRAKYDLYNNYVNLAKEVEKAGAHIIAKDMA 726

QY 734 GLLRPAAYTKLVATLALREFDLPHVHTHTAGGOLATYFAAAQAGADAVDASAPISGTT 793

Db 727 GLLKPAARVLFKALREEDLPIHFHTHTDTSIGISAATVLAADGAVDVVDAAAMDALSGNT 786

QY 794 SQPSLSAIVAAFAHTRDGTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIP 853

Db 787 SQPCLSGIVEALHGSERDGLDPLRRISFYWEAVRHOYAAEFESDLKGPASEVYLHEMP 846

QY 854 GGQSLNLRQAATLADLAFELIEDNAAVNEMLRPTKVTPSSKVVYGDIALHLHAGVGD 913

Db 847 GGOFTNLKQEARSLGLETRHWEVAQAYADVRNMGDIVKVTTPSKVVGDMLAMVAAQDLT 906

QY 914 PADFAADPOKYDIPDSVIAFLRGELNPPGWPPELIRALEGRSECKAPLTVPEE--E 971

Db 907 VADVENPAKDIAFPDSVSWMMRGDLGQPPSGWPEALQKKVL----KDEKPTVRPGSLP 962

QY 972 QAHLDA-----DDSKERNRS----LNRLLPKPTSEFEHRRRFGNTSALDDREFFYGLV 1022

Db 963 AADLDAERKSFEDSVGKRLSQDQEFASALMYPKVFTDYATAHETGYPTSVLPTVYFYGLK 1022

QY 1023 EGRETILIRLPDVRTPLLRDLDAISEPDDKGMNRVNVANVGDIRPMRVDRDRSVESVTAAE 1082

Db 1023 PEEEVFDVLERGKTLVIVN-QAMSETDEKGMVTVPFELNGQPRIKVPNRAKASGVRR 1081

QY 1083 KADSSNKGHVAAPFAGVV-TVTVAEGDEVKAGDAVAIIEMKMEATITASVDGKIERRVV 1141

Db 1082 KVEAGNDKQVGAPMPCVISTVAVVAGQKVTQGDVLLSIEAMKMETAIHAERDGTIAEVLV 1141

QY 1142 PAATKVEGGDLIVV 1155

Db 1142 RPGEQIDAKDILLIV 1155

RESULT 14

T20346

pyruvate carboxylase (EC 6.4.1.1) D2023.2 [similarity] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Aug-2000

C:Accession: T20346

R:Kershaw, J.

submitted to the EMBL Data Library, October 1996

A:Reference number: Z19260

A:Accession: T20346

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-1175 <WIL>

A:Cross-references: EMBL:Z81052; PIDN:CAB02872.1; GSPDB:GN00023; CESP:D2023.2

A:Experimental source: clone D2023

C:Genetics:

A:Gene: CESP:D2023.2

A:Map position: 5

A:Introns: 34/3; 103/3; 246/1; 451/3; 567/2; 821/1; 1045/3; 1126/3

C:Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binding

C:Keywords: ligase

F:1140/Binding site: biotin (lys) (covalent) #status predicted

Query Match 42.0%; Score 2461; DB 2; Length 1175;

Best Local Similarity 45.1%; Pred. No. 2.3e-127;

Matches 522; Conservative 201; Mismatches 389; Indels 46; Gaps 14;

QY 29 FKKILVANRGEIARAFRAALETGAATVAIYPRDRSGFHRSFASEAVRITGEGSPVAY 88

Db 32 FNKVMVANRGEIARVFRALTELNLKTSVAIYAEQDKNSMHLKADAEYLVKGLPPVAAV 91

QY 89 LDIDEIIGAARKVKADAIYPGYFLSENQAALRECAENGIPTGTPVEVLDTGDKSRV 148

Db 92 LTIDQIETIETALHNIDAIHPGYGFLSERSDFAAACAGIVFICGSPDVNARMGDKVAAR 151

QY	149	TAAKAGLPVLAEST-PSKNIDDDIVKSAEGQYPIFYKAVAGGGGGRGMFVSSPDELRLK	207
Db	152	QAAIEAGVQVVGPGPTTTTDAEAVEFAKQYGTPIILKAAYGGGGRGIRRVLDLEEVEEA	211
QY	208	ATEASREAAAFGDSGVYVERAVINPOHIEVQIILGDRGTGEVHLYERDCSLQRHQRKVV	267
Db	212	FRSYSEAAAFGDSGLFVEKFEVERPHIEVQLLGDHGNVHLYERDCSVQRHQRKVV	271
QY	268	IAPAQLDPELDRICADAVKFCRSIGVGAGTVEFLVDEKGNHVFLEMPRIQVEHTVT	327
Db	272	IAPALPEGVREKILADALRLARHVGYQAGTVEFLVDQKGNFYFEVNRARQVEHTVT	331
QY	328	EEVTEVDLVKAQMLAAGATLKLGTQDKTHGAALQCRITTEDPNNNGFRPDGTGITA	387
Db	332	BEITGVDLVQAQIRIAEGSKLDDLKLSQETIQTGSAIQCRVTTEDPAKQFQDSGRIEV	391
QY	388	YRSPGGAGVRLDGAALQGE-ITAFPSDMLVKTCRSDSETAVARAQALAEFTVSGVA	446
Db	392	FRSEGMGIRLDSASAFAGSVISPHYSDLMAVKVITASARNHPNAAKMIRALKFRINGVK	451
QY	447	TNIGFLRALLREDEFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLVADVTYVKNP	502
Db	452	TNIPFLNLVLRQPSFLDASVDYTFIDEHPLOFPQKPSQNRQAQKLLNLYLGEVKNPPTPL	511
QY	503	-HGVRPKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAFAFARDLREQDALAVDTTFRDA	561
Db	512	ATDLKPAVVSPIPIYIP--AGAKPPTGLRDVLVQVGTEFAKEVRSRPGCMITDTTFRDA	569
QY	562	HQSLATRVRSFALKPAEAAVAKLTPELLSVFAMGGGATYDVAMRFLPEDPDRDLRELA	621
Db	570	HQSLATRVRTYDMAAISPFAVQFNGFLFLENWGGATFDVSMRFLCECPWERLQTLRLK	629
QY	622	MPNYNIQMLLRNTGVTPYPSVCFRAVKEAASSGVDIFRIFDALNDVYSQMRPALDAV	681
Db	630	IPNIPFQCLLRGANAGYSNYPDNVYKFCVELAKVNGMDVRFVDSLNYLPNLLVGMVAV	689
QY	682	LETNTAAEAVAMAYSGDLSDDPNKLYTLDYLYKMAEIVKSGAHILAIKDMAGLLRPAV	741
Db	690	GKAG-GVVEAAIAVTGDDVTDKSRDKYLDKYYLNLADQLVKAQAHILSKDMAGVLRPEA	748
QY	742	TKLVTALRRP-DLPVHVHTHTAGGOLATYFAAAQAGADAVDQASAPLSCGTSQPSLSA	800
Db	749	KLLICALRDKFPDPIHVHTHTDSAGVAAMLECAKAGADVDAVDAVDSMGMTSQPSMGA	808
QY	801	IVAAFAHTRRTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGR--VYRHEIPGQOLS	858
Db	809	IVASLQGTGKHTGLSDDISKYSAYWESTKQLYAPFECATMKSGNADVYKHEIPGQYTT	868
QY	859	NLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPTSPKSVGDLALHLVAG-----V	912
Db	869	NLQQAFLSLGPGQFDEVKRYREANLVGLDIIKVTPESSKIVGDLAQFMVQNNLTRTLV	928
QY	913	DPADFAADPOKQYIPDSVIAFLRGELGNPPGWPPEPLRTRALEGRS--EGKAPLTVPEE	970
Db	929	DRAD-----DLSPKSVSDPMQGNVGPYPGPFPEPLRTKVLGRKPKVDGR-----PGE	976
QY	971	EQALHDADDK-----ERNRSLNRLLPKPTPEEFLEHRRRFGNTSALDDREFFY	1019
Db	977	NAKPVLDLAVKVELEEKHGRTLSSEDMYSMEPTVDFETFRQQYGPVDKLPTRLFLT	1036
QY	1020	GLVEGRETILRLPDVRYPLLVRLDAISEPDDKGMNRNVVANNVQIRPMRYRDRSVSVTA	1079
Db	1037	GLETAEEVDVEIESKT-LAIQLLAELKLNKRGEREVFFDLNGOMRSIFVVDKEASKEIV	1095
QY	1080	TAEKADSSNKGHVAAPFAG-VVTVTVAGDEVKAGDAVAIIIEAMKEATTITASVDGKIER	1138
Db	1096	TRPALPCVRGHICAPMPGDVLELKIKEGDKVYKKQPLFVLISAKMDEMWDISPIAGTVKA	1155
QY	1139	VVVAATKVEGGDLIVVV 1156	
Db	1156	IHAPQGTKCSAGDLVVEV 1173	

RESULT 15
QYBYB
pyruvate carboxylase (EC 6.4.1.1) 1 [validated] - yeast (Saccharomyces cerevisiae)
NAlternate names: protein G3428; protein YGL062w; pyruvic carboxylase
C:Species: Saccharomyces cerevisiae
C:Date: 31-Dec-1991 #sequence_revision 19-Jul-1996 #text_change 01-Feb-2002
C:Accession: S64066; A29233; S05722; A29722
R:Feuermann, M.; Potier, S.; Souciet, J.L.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64044
A:Accession: S64066
A:Molecule type: DNA
A:Residues: 1-1178 <FEU>
A:Cross-references: EMBL:272584; NID:gl322565; PIDN:CAA96765.1; PID:gl322566; GSPDB:G
A:Experimental source: strain S288C
R:Lim, F.; Morris, C.P.; Occhiodoro, F.; Wallace, J.C.
J. Biol. Chem. 263, 11493-11497, 1988
A:Title: Sequence and domain structure of yeast pyruvate carboxylase.
A:Reference number: A92662; MUID:88298805; PMID:3042770
A:Accession: A29233
A:Molecule type: DNA
A:Residues: 1-461, 'G', 463-492, 'D', 494-594, 'A', 596-618, 'O', 620-663, 'S', 665-771, 'R', 773
A:Cross-references: EMBL:J03889; NID:gi72101; PIDN:AAA34843.1; PID:gi72102
R:Morris, C.P.; Lim, F.; Wallace, J.C.
Biochem. Biophys. Res. Commun. 145, 390-396, 1987
A:Title: Yeast pyruvate carboxylase: gene isolation.
A:Reference number: S05760; MUID:87241529; PMID:3036126
A:Accession: S05760
A:Molecule type: DNA
A:Residues: 1003-1178 <MOR>
A:Cross-references: EMBL:J03889
A:Accession: A29722
A:Molecule type: protein
A:Residues: 1124-1149 <MOR2>
C:Genetics:
A:Gene: SGD:PYC1; PYV: MIPS:YGL062w
A:Cross-references: SGD:S0003030; MIPS:YGL062w
A:Map position: 7L
C:Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-biotin-binding; biotin binding; gluconeogenesis; homotetramer; ligase; zinc
F:21,478/Domain: Biotin carboxylase homology <BCH>
F:157-33/Domain: ATP/bicarbonate binding #status predicted <ATBI>
F:353-468/Domain: ATP/bicarbonate binding #status predicted <ATB2>
F:569-908/Domain: pyruvate binding #status predicted <PYR>
F:1096-1169/Domain: lipoyl/biotin-binding homology <LPB>
F:1135/Binding site: biotin (Lys) (covalent) #status experimental

Query Match 41.9%; Score 2457; DB 1; Length 1178;
Best Local Similarity 46.6%; Pred. No. 3.8e-127;
Matches 540; Conservative 175; Mismatches 400; Indels 44; Gaps 16;

QY	31	KILVANRGETIAVRAALETGAATVAIYPRDRGSRFHSFASAEVRIQTEG--SPVKAY	88
Db	21	KILVANRGETIPRIPTAHLSNQTVAIYSHEDRLSTHKOKADEAYVIGEVQYTPVGA	80
QY	89	LQDEITGAIAKVKADAIYGYGFLSENAOLARECAENGITFIPTEVLDLTGDKSRV	148
Db	81	LAIDEIISIAQKHQVDFIHPGYGFLSENSEFADKVKVAGITWIGPPAEVIDSVGDKVSAR	140
QY	149	TAAKAGLPVLAEST-PSKNIDDIVKSAEGQYPIFYKAVAGGGGGRGMFVSSPDELRLK	207
Db	141	NLAARANVTVPCTPCPIETVEALDFVNEYGYPIIKAAGGGRGMVREGDDVADA	200
QY	208	ATEASREAAAFGDSGVYVERAVINPOHIEVQIILGDRGTGEVHLYERDCSLQRHQRKVV	267
Db	201	FORATSEATAFNGTCFVERFLDKPHIEVOLLADNHNVHLYERDCSVQRHQRKVV	260
QY	268	IAPAQLDPELDRICADAVKFCRSIGVGAGTVEFLVDEKGNHVFLEMPRIQVEHTVT	327
Db	261	VAPAKTLPREVREDAITLTDVAKLAKKEGYNAGTAEBFLVDNQNHRHFIEINPRIQVEHT	320
QY	328	EEVTEVDLVKAQMLAAGATLKLGTQDKTHGAALQCRITTEDPNNNGFRPDGTGITA	387

Db 321 EBITGIDIVAAQIAAGASLPQLGLFQDKITTRGFATQICRITTEDPAKNFQPDGTGRIEV 380
QY 388 YRSPGGAGVRLDGA-AQLGGEITAHFDSMLVKMTCRGSDFFETAVARAORALAEFTVSGVA 446
Db 381 YRSAGNGVRLDGGNAYAGTIIISPHYDSMLVWCSCSGSTYEVRRMRIRALIEFRIRGVK 440
QY 447 TNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEOQRIILDYLADEVN---KP 502
Db 441 TNIPELLTLLTNPVFIEGTYYWTFIDDTPLQFQMYSSQNRAQKLLHYLADAVANGSSIIG 500
QY 503 HGVPRKDVAAPIDKLPNTKDL-----PLPRGSRDLKQLGPAAPARDLREODALA 552
Db 501 QTGLPKLKNP--SYPHLHDAQGNVINTKSAPPSGWQVLLKEKPAEFARQVROFNGLT 558
QY 553 VTDITFRDAHQSLLATRYRVSFALPKAAEAVAKLTPELLSVEAMGGATYDVAMRFLFEDPW 612
Db 559 LMDTTRDAHQSLLATRYRVTHTDLATIAPTTAHALACRFALECGGATFDVAMRFLHEDPW 618
QY 613 DRDELREAMPNVNIOMLLGRNTYGYTPYDSCVCAFVKEAASSGVDFIRIFDALNDVS 672
Db 619 ERLRLRSILVPNIPEQMLLRGANGVAYSSLPDNAIDHFVKQAKDNGVDIFRVFDALNDLE 678
QY 673 QMRPAIDAVLENTAVAEVAMAYSGLSDPNNEKLYTLOYLLKMAEBEIVKSGAHILAIKDM 732
Db 679 QLKVGVDVAKKAG-GVVEATVCFSGDMLQPGKK-YNLDYILEIAEKIVQMGTIILGIKDM 736
QY 733 AGLLRPAAVTKLVATLAREF-DLPVHVHTHTAGQLATYFAAAQAGADAVDGSAPLSG 791
Db 737 AGTMPAAAKLLIGSLRAKYPDLPPIHVHTHDSAGTAVASMTACALAGADVVDVAINMSG 796
QY 792 TTSQPSLSAIVAAFAHTRDRTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHE 851
Db 797 LTSQPSINALLASL-EGNIDTGINVHVRELDAYWAEMRLLYSCFEADLKGPDPVEYQHE 855
QY 852 IPGGQLSNLRQAOTALGLADRELELIEDNYAAVNEMLGRPTKVTSPSKVVGDLALHLVGAG 911
Db 856 IPGGQLTNLLFOAQQLGLGEQWAEKRAYREANYLLGDIVKVTPTSKVVGDLAQFVSNK 915
QY 912 VDPADFAADPOKYDIPDSVIAFLRGLGNPPGPGWPEPLRTRALEGRSEKAPLTVPEEE 971
Db 916 LTSDDVRLANSLDFPDSVMDFEGLIGQPYGCFPEPFRSDVLRNK---RRKLCRPGLE 972
QY 972 QAHLDADDKERNLSNRL-----LFPKPTEEFLEHRRFRFGNTSALDDREFFY 1019
Db 973 LEPPDLE--KIREDLQNRFGDVDECDVASYNMYPRIYEDFKMRETYGDLVLPTRSFLS 1030
QY 1020 GLVEGRETLIRLPDVRTPLLVRLDAISEPDDK-GMRNVVANVNGOIRPMRVDRDSVESVT 1078
Db 1031 PLETDEEIEVWIEQKLT-LIILQAVGDLNKKTGEREYFDLNGEMRKIRVADRQKQVET 1089
QY 1079 ATAERKADSSNKGHVAAPFAGV-VTVTVAGDEVKAGDAVAIIEAMKMEATITASVDGKIE 1137
Db 1090 VTKSRADMDPLHIGAPMAGVIVEVKVHKSGLIKGQPVAVLSAMKMEMIISPSDGGQVK 1149
QY 1138 RVVPAATKVEGGDLIVVV 1156
Db 1150 EVFVSDGENVDSDLVLVL 1168

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2003, 18:55:00 ; Search time 16 Seconds
(without alignments)
2999.257 Million cell updates/sec

Title: US-09-974-973-2
Perfect score: 5865
Sequence: 1 MTATLGLLKGITLVT.....RVVPAATKVEGDLIVVVS 1157

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Query Length	ID	Description
1	2550.5	43.5	1178	1 PYC_MOUSE	Q05920 mus musculus
2	2534.5	43.2	1178	1 PYC_RAT	P52873 rattus norv
3	2531.5	43.2	1178	1 PYC_HUMAN	P11498 homo sapien
4	2457	41.9	1178	1 PYC1_YEAST	P11154 saccharomyc
5	2436	41.5	1180	1 PYC2_YEAST	P32327 saccharomyc
6	2393	40.8	1189	1 PYC_PICPA	P78992 pichia past
7	1079	18.4	501	1 PYCA_METJA	Q58626 methanococ
8	1032.5	17.6	447	1 ACCC_ANASP	Q06862 anabaena sp
9	1020	17.4	506	1 PYCA_ARCFU	Q30019 archaeglob
10	996	17.0	491	1 PYCA_METTH	O27939 methanobact
11	950	16.2	448	1 ACCC_HAEIN	P43873 haemophilus
12	940	16.0	449	1 ACCC_ECOLI	P24182 escherichia
13	938	16.0	449	1 ACCC_ECO57	Q8x366 escherichia
14	921	15.7	449	1 ACCC_PSEAE	P37798 pseudomonas
15	897	15.3	725	1 MCCA_HUMAN	Q967q3 homo sapien
16	894	15.2	717	1 MCCA_MOUSE	Q99mr8 mus musculus
17	891.5	15.2	654	1 BCCA_MYCTU	P46401 mycobacteri
18	875	14.9	448	1 ACCC_BACSU	P49787 bacillus su
19	874.5	14.9	567	1 PYCB_METJA	Q58628 methanococ
20	869.5	14.8	598	1 BCCA_MYCLE	P46392 mycobacteri
21	869	14.8	703	1 PCCA_HUMAN	P05165 homo sapien
22	865	14.7	731	1 MCCA_SOYBN	Q42777 glycine max
23	862	14.7	734	1 MCCA_ARATH	Q42523 arabidopsis
24	842.5	14.4	704	1 PCCA_RAT	P14882 rattus norv
25	837.5	14.3	1835	1 DURL_YEAST	P32528 saccharomyc
26	828	14.1	568	1 PYCB_METTH	O27179 methanobact
27	806	13.7	590	1 DCOA_SALTY	Q03030 salmonella
28	801.5	13.7	595	1 DCOA_KLEPN	P13187 klebsiella
29	620.5	10.6	2345	1 COAL_RAT	P11497 rattus norv
30	617.5	10.5	2324	1 COAC_CHICK	P11029 gallus gall
31	617.5	10.5	2346	1 COAL_HUMAN	Q13085 homo sapien
32	616.5	10.5	2346	1 COAL_BOVIN	Q9ttt3 bos taurus
33	607.5	10.4	2346	1 COAL_SHEEP	Q28559 ovis aries

ALIGNMENTS

RESULT 1

ID	PYC_MOUSE	STANDARD;	PRT;	1178 AA.
AC	Q05920:			
DT	01-FEB-1994 (Rel. 28, Created)			
DT	01-FEB-1994 (Rel. 28, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Pyruvate carboxylase, mitochondrial precursor (EC 6.4.1.1) (Pyruvic			
DE	carboxylase) (PCB).			
GN	PC OR PCX.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Adipocyte;			
RX	MEDLINE=93189578; PubMed=8446588;			
RA	Zhang J., Xia W.L., Brew K., Ahmad F.;			
RT	"Adipose pyruvate carboxylase: amino acid sequence and domain			
RT	structure deduced from cDNA sequencing."			
RL	Proc. Natl. Acad. Sci. U.S.A. 90:1766-1770(1993).			
CC	-!- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,			
CC	INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY			
CC	ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE			
CC	CARBOXYL GROUP TO PYRUVATE IN THE SECOND. CATALYZES IN A TISSUE			
CC	SPECIFIC MANNER, THE INITIAL REACTIONS OF GLUCOSE (LIVER, KIDNEY)			
CC	AND LIPID (ADIPOSE TISSUE, LIVER, BRAIN) SYNTHESIS FROM PYRUVATE.			
CC	-!- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3)(-) -> ADP + phosphate +			
CC	oxaloacetate.			
CC	-!- COFACTOR: BIOTIN AND MANGANESE.			
CC	-!- PATHWAY: GLUCONEOGENESIS AND LIPOGENESIS.			
CC	-!- SUBUNIT: HOMOTETRAMER.			
CC	-!- SUBCELLULAR LOCATION: Mitochondrial matrix.			
CC	-!- TISSUE SPECIFICITY: LIVER, KIDNEY, ADIPOSE TISSUE, LIVER			
CC	AND BRAIN.			
CC	-!- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOAMIDE TRANSFERASES			
CC	AND CARBAMYL PHOSPHATE SYNTHETASES.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; L09192; AAA39737.1; -			
DR	PIR; A47255; A47255.			
DR	HSSP; P24182; 1BNC.			
DR	SWISS-2DPAGE; Q05920; MOUSE.			
DR	MGD; MGI:97520; PCX.			
DR	InterPro; IPR001882; PCX.			
DR	InterPro; IPR000089; Biotin_lipoyl.			
DR	InterPro; IPR000901; CPSase.			
DR	InterPro; IPR000891; HMGL-like.			

P32874 saccharomyc
O00763 homo sapien
Q00955 saccharomyc
P78820 saccharomyc
O28994 archaeglob
O27077 methanobact
Q8xa38 escherichia
Q9wz27 thermotoga
P00968 escherichia
P14846 salmonella
Q82917 salmonella
P03965 saccharomyc

34 570.5 9.7 2273 1 HFA1_YEAST
35 560.5 9.6 2483 1 COA2_HUMAN
36 547.5 9.3 2233 1 COAC_YEAST
37 519 8.8 2280 1 COAC_SCHPO
38 267.5 4.6 1076 1 CARB_ARCFU
39 266 4.5 1060 1 CARB_METTH
40 255 4.3 1072 1 CARB_ECO57
41 254.5 4.3 1099 1 CARB_THEMEA
42 252 4.3 1072 1 CARB_ECOLI
43 248.5 4.2 1074 1 CARB_SALTY
44 245.5 4.2 1074 1 CARB_SALTY
45 244 4.2 1118 1 CARB_YEAST

DR InterPro: IPR003379; Pyc_OADA.
 DR Pfam: PF00289; Cpsase_L-chain; 1.
 DR Pfam: PF00364; biotin_lipoyl; 1.
 DR Pfam: PF00682; HMGU-like; 1.
 DR Pfam: PF02436; Pyc_OADA; 1.
 DR Pfam: PF02785; Biotin_carb_C; 1.
 DR Pfam: PF02786; Cpsase_L_D2; 1.
 DR Pfam: PF02786; Cpsase_L_D2; 1.
 DR TIGRFAMs: TIGR01235; pyruv_carbox; 1.
 DR PROSITE: PS00188; BIOTIN; 1.
 DR Ligase; Multifunctional enzyme; Biotin; Manganese; Gluconeogenesis;
 DR ATP-binding; Mitochondrion; Lipid synthesis; Transient peptide.
 DR TRANSIT 1 20 MITOCHONDRION (POTENTIAL).
 DR CHAIN 21 1178 PYRUVATE CARBOXYLASE.
 DR DOMAIN 21 549 BIOTIN CARBOXYLASE (BY SIMILARITY).
 DR DOMAIN 550 1000 CARBOXYLTRANSFERASE (BY SIMILARITY).
 DR DOMAIN 1096 1178 BIOTIN CARBOXYL CARRIER PROTEIN
 DR (BY SIMILARITY).
 DR NP_BIND 198 203 ATP (BY SIMILARITY).
 DR ACT_SITE 328 328 BY SIMILARITY.
 DR BINDING 1144 1144 BIOTIN (BY SIMILARITY).
 DR SEQUENCE 1178 AA; 129684 MW; 14CEA0F9DA8B8127 CRC64;

Query Match 43.5%; Score 2550.5; DB 1; Length 1178;
 Best Local Similarity 46.1%; Pred. No. 2, 5e-129;
 Matches 543; Conservative 192; Mismatches 407; Indels 35; Gaps 13;

QY 7 GGLLKGITLVSTHTSTLPA-----FKILVANGETIAVRAFAALETGAA 54
 DB 9 GGLRLGV-----RRSSAPVSPNRRLEYKPIKVVANGETIAVRAFAALETGIR 62

QY 55 TVAIYPREDGSRFHSRASEAVRIGTSGPVKAYLDIDEIIGAARKVADAIYVGGYGLS 114
 DB 63 TVAVISEDTGMHQRKADAEYLIGRLAPVQAYLHPDIKIVAKENGVDVHPGYGLS 132

QY 115 ENAOLARECAENITFIGTPEVLDTLGDKSRAVTAAKAGLPVL-AESTPSKNIDIVK 173
 DB 123 ERAFAQACDAGVRFTGPSSEVYKMGDKVEARAIAAAGVVPVPGTDSPISSLHEAE 182

QY 174 SAEGQTPYFKAVAGGGGGRMREVPSPDELRLKLTAEASREAAFGGSGVYVERAVINP 233
 DB 183 FSNFTGFPPIPKAAYGGGGRMVRVHSEYLEENYTRAYSEALAAFGALFVEKTEKP 242

QY 234 QHIEVQILGDRTEGVVHLRYDCSLQRHQKRVIEAPAOHLDPQLRDRICADAVKFCRSI 293
 DB 243 RHIEVQILGDQYGNILHYLRDCSIQRHQKRVIEAPATHLPQLRSLTSDSVKLAKOV 302

QY 294 GYQAGTVEFLVDEKGNHVFTEMPRIQVEHTVTEVTEVDLVKAQMLAAGATLKLGL 353
 DB 303 GYENAGTVEFLVDEKGNHVFTEMPRIQVEHTVTEVTEVDLVKAQMLAAGATLKLGL 362

QY 354 TQDKIKTHGAALOCRTTEDPNNGFRPDTGTITAYRSPGGAGVRLDGAAL-GGEITAHF 412
 DB 363 ROENIRINGCAIOCRVTTEDPARSFQDPTGRIEVRSGEGMIRLDNASAFQAVLSPHY 422

QY 413 DMLVKMTCRGSDPETAVARAQALAEFTVSGVATNIGFLRLALREDEFTSKRIATGFIG 472
 DB 423 DMLVKMTCRGSDPETAVARAQALAEFTVSGVATNIGFLRLALREDEFTSKRIATGFIG 482

QY 473 DHPHLLQAPADDEGRITLDYADVTNKGVPKQVA-APID-KLPNTKDLPLRGRS 530
 DB 483 ENPELFQALPAQNAQKLLHYLGHVWNGTTPPIPVNWSPPDPAVVPVPGPPAGFR 542

QY 531 DRKLQGPAPARDLRODALVDTTFRDAHQSLTARVRSPALPAEAAVAKLTPELL 590
 DB 543 DILLREGPEGARAVRNHQGLLMDTFRDAHQSLTARVRHDLKRIADYVAHNKLF 602

QY 591 SVEAGGATYDVAMRFLFEDPWRDLDEAMPNVNQMILLRGNTVGYTPYDSCVRAF 650
 DB 603 SMENMGATYDVAMRFLFEDPWRDLDEAMPNVNQMILLRGNTVGYTPYDSCVRAF 662

QY 651 VKEAASSGVDFIRFDALNDVDSQMRPAIDVLETNTTAVAEVAMAYSGDSDPNKLYTLD 710
 DB 663 CEVAKENGMDVFRVDFSLNLPNMLLMEAGSAG-GVVEAAISYTGADVADPSRTRYSE 721

QY 711 YLKMMAEIVKGAHILAIKDMAGLRLPAAVTKLVTAALREF-DLPVHVHTHTDTAGQOLA 769
 DB 722 YNGLAEELVRAGTHILCIKDMAGLRLKPAACMLVSSLRDRFPDLPLHHTHTDTAGVA 781

QY 770 TYFAAQAAGADAVDGAAPLSGTTTSPQSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAV 829
 DB 782 AMLACAQAAGADVVDVAVDSMSGMTSPQSMGALVACTKGTPLDTEVPFLRVFDYSEYEGA 841

QY 830 RGLYLPFESGTPGPTGR--VYRHEIPGGQLSNLRAQATGALGLADLFELEDNYAAVNEML 887
 DB 842 RGLYAAFDCTATMKSGNSDVIENIEPGGYTNLHFQAHSMGLSGSKFVKKAYVEANQML 901

QY 888 GRPTKVTPTSSKVVGDIALHLVAGADPADADPOKYDIPDSVIAFLRGELGNPPGGWPE 947
 DB 902 GDILKVTPTSSKIVGDIAQPMVQNGLSRAEAAEAQAEELSPRSVVEFLOGYIGIPHGPE 961

QY 948 PLRTRAL-----EGRSECKAPLTVPEPEOAHLD--DDSKERNSLNRLFPKTEEFL 1000
 DB 962 PFRSKVLKDLPRIEGPGASLPNLKLEKDLIDRHGEVTPEDVLSAAMYDPVPAQFK 1021

QY 1001 EHRREFGNTSALDDREFFYGLVEGRETILRLPDVTPTLVRLDAISEPDDKGRNVVAV 1060
 DB 1022 DTAFTGPLDSNTRFLFLOGPKIAEEFEVELERGKT-LHIKALAVSDLNRAQOVFFEL 1080

QY 1061 NGQIRPMVRDRSVESVYATAEKADSSNKGHVAAPAG-VVTVTVAEQGVKAGDAVAII 1119
 DB 1081 NGQLRSILVKDTQAMKEMHFHFKALKDKVQIGAPMPGKVIDIKVAAGDKVAKGQPLCVL 1140

QY 1120 EAMKMEATTITASVDGKIERYVVPVPAATKVEGGDLIVVV 1156
 DB 1141 SAMKMETVTVSPMEGTIRKVVHTKDMTLEGDDLILEI 1177

RESULT 2
 PYC_RAT
 ID PYC_RAT STANDARD; PRT: 1178 AA.
 AC P52873; Q64555;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Pyruvate carboxylase, mitochondrial precursor (EC 6.4.1.1) (Pyruvic
 DE carboxylase) (PCB).
 GN PC.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1] SEQUENCE FROM N.A.
 RP TISSUE=Liver;
 RX MEDLINE=96096548; PubMed=8522203;
 RA Lehn D.A., Moran S.M., Macdonald M.J.;
 RT "The sequence of the rat pyruvate carboxylase-encoding cDNA.";
 RL Gene 165:331-332(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar;
 RX MEDLINE=96257760; PubMed=8687410;
 RA Jitrapakdee S., Booker G.W., Cassidy A.I., Wallace J.C.;
 RT "Cloning, sequencing and expression of rat liver pyruvate
 RT carboxylase.";
 RL Biochem. J. 316:631-637(1996).
 CC -!- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
 CC INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
 CC ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE
 CC CARBOXYL GROUP TO PYRUVATE IN THE SECOND. CATALYZES IN A TISSUE
 CC SPECIFIC MANNER, THE INITIAL REACTIONS OF GLUCOSE (LIVER, KIDNEY)
 CC AND LIPID (ADIPOSE TISSUE, LIVER, BRAIN) SYNTHESIS FROM PYRUVATE +
 CC -!- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3)(-) = ADP + phosphate +
 CC oxaloacetate.
 CC -!- COFACTOR: BIOTIN AND MANGANESE (BY SIMILARITY).
 CC -!- PATHWAY: GLUCONEOGENESIS AND LIPOGENESIS.

CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
 CC -1- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOAMIDE TRANSFERASES
 CC AND CARBAMYL PHOSPHATE SYNTHETASES.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: U32314; AAC96256.1; -;
 CC EMBL: U36585; AAC52668.1; -;
 CC HSSP: P24182; 1BNC.
 DR InterPro: IPR001892; Biotin_attach.
 DR InterPro: IPR000089; Biotin_lipoyl.
 DR InterPro: IPR000901; CPSase.
 DR InterPro: IPR000891; HMG-L-like.
 DR InterPro: IPR003379; PYC_OADA.
 DR Pfam: PF00289; CPSase_L_chain; 1.
 DR Pfam: PF00364; biotin_lipoyl; 1.
 DR Pfam: PF00682; HMG-L-like; 1.
 DR Pfam: PF02436; PYC_OADA; 1.
 DR Pfam: PF02785; Biotin_carb_C; 1.
 DR Pfam: PF02786; CPSase_L_D2; 1.
 DR TIGRFAMS: TIGR01235; pyruv. carbox; 1.
 DR PROSITE: PS00188; BIOTIN; 1.
 DR Ligase; Multifunctional enzyme; Biotin; Manganese; Gluconeogenesis;
 KW ATP-binding; Mitochondrion; Lipid synthesis; Transist peptide.
 KW TRANSIT 1 20 MITOCHONDRION (POTENTIAL).
 FT CHAIN 21 1178 PYRUVATE CARBOXYLASE.
 FT DOMAIN 21 549 BIOTIN CARBOXYLASE (BY SIMILARITY).
 FT DOMAIN 550 1000 CARBOXYLTRANSFERASE (BY SIMILARITY).
 FT DOMAIN 1096 1178 BIOTIN CARBOXYL CARRIER PROTEIN
 FT (BY SIMILARITY).
 FT NP_BIND 198 203 ATP (BY SIMILARITY).
 FT ACT_SITE 328 328 BY SIMILARITY.
 FT BINDING 1144 1144 BIOTIN (BY SIMILARITY).
 FT CONFLICT 222 222 P -> S (IN REF. 2).
 FT CONFLICT 866 866 D -> I (IN REF. 2).
 FT CONFLICT 977 977 G -> R (IN REF. 2).
 SQ SEQUENCE 1178 AA; 129689 MW; 8E5FA19BC132A8DD CRC64;
 Query Match 43.2%; Score 2534.5; DB 1; Length 1178;
 Best Local Similarity 46.2%; Pred. No. 1.8e-128;
 Matches 544; Conservative 188; Mismatches 409; Indels 37; Gaps 14;
 QY 7 GGLLAGIITLVSTHTSSTLP-----FKILVANRGEIAVRAALETGAA 54
 DB 9 GGLRLGV-----RSSTAPASPNVRRLLEYKPIKVVANRGEIAVRACTELGIR 62
 QY 55 TVAIYPREDRGSHRSFASBAVRIGTEGSPVKAYLDEIIGAIAKVKADAIYPGYGFLS 114
 DB 63 TVAVISEQDTGOMHROKADAYLIGRLAPVQAYLHPDIKIVAKENGVDVHPGYGFLS 122
 QY 115 ENAQLARECAENGITFTGPTPEVLDTGDKSRATVAAKAGLPVL-AESTPSKNIDIVK 173
 DB 123 ERADFAQACODAGVRFTIGPSPEVVRKMGDKVEARATAAAGVPVPGTNSPINSLHEAHE 182
 QY 174 SABGQOTYPIFKAVAGGGRGMFVSPDELRLKLAETASREAAAFDGDGVSYYVERAVINP 233
 DB 183 FSNTYGFPIIFKAYGGGGRGMVHVSYEELEENYTRAYPEALAAFGNGALFVEKTEKP 242
 QY 234 QHTEVILGDRTEGVHLYERDCSLQRRHQKVVVEIAPAQHLDELDRICADAVKFCRSI 293
 DB 243 RHIEVQILGDOYGNILHLYERDCSIQRRHQKVVVEIAPATHLDPQLRSRLTSDSVKLAKQV 302
 QY 294 GYOGACTVPELVDEKGNHVFIEKNPRIOVHTTVEETVEDVLYKAQNRMLAAGATLLELGL 353
 DB 303 GYENAGTVEFLVDKHKHGYEIVNSRLQVEHTTVEETIDVDLVAHQIHVSEGRSLPDLGL 362

QY 354 TQDIKIKTHGAALOCRIITTPNNGFRPDGTGTITAYRSPGGAGVRLDGAAL- GGEITAHF 412
 DB 363 RQENIRINGCAIQCRVTEDPARSFQDPTGRIEVRSEGGNGIRLDNASAFQAGAVISPHY 422
 QY 413 DMLVVKMTCRGSFETVARAQRALAEFTVSGVATNIGFLRALLREBDEFTSKRIATGFIG 472
 DB 423 DSSLVKVIAHGKHDPHTAATKMSRALAEFRVGVKTNIPFLQNVLNNOQFLAGIVDTOFID 482
 QY 473 DHPHLLQAPADDQGRILYDLADVTYVKNKHGVRPKDVA-APIDKL-PNIKDLPLPGSR 530
 DB 483 ENELFQLRPAQNRAQKLLHYLGHVWYNGVPTPIPVKVSFSPDPIVPPVIGIPPPAGFR 542
 QY 531 DRLLKQLGPAFAFARDLRDQDALAVTDITFRDAHQSLLATRVRSFALKPAAEAVALKTPELL 590
 DB 543 DILLREGPEGFARAVRNHOGLLMDITFRDAHQSLLATRVTHDKKTIAPYVAHFNFLF 602
 QY 591 SVEAWGATYDVAMRFLFEDPWRDLDELREAMPNWNTOMLLRGNTVGYTPYDPSVCRAP 650
 DB 603 SIENWGGATFDVAMRFLYECPRRLQELRLIPNIPFQMLLRGANAVGYTNPONVVFKE 662
 QY 651 VKEAASSGVDFIERFDALNDVSOMRPAIDAVLENTTAAVEAMAYSGDLSDPNEKLYTLD 710
 DB 663 CEVAKENGMDVFRIFDLSNLYLPNMLLGEAAGSAG-GVVEAAISYTGVDADPSTKYSLE 721
 QY 711 YYLKMAEEIVKGAHILAIKDMAGLLRPAAVTKLVLTALRREF-DLPVHVHTHDTAGQLA 769
 DB 722 YMWGLAEELVRAGTHILCIKDMAGLLKPACTMLVSSLRDFPDLPLHIHTHDTSGSVA 781
 QY 770 TYFAAQAGADAVDGSAPLSGTTSQPSLSAIVAAFAHTRDRTGLSLEAVSDLEPYWEAV 829
 DB 782 AMLACAQAGADVVDVAVDSMSGMTSQPSMGALVACTKGTPLDTEVPLERFVDFSEYWEA 841
 QY 830 RGLYLPFESCTPGPTGR--VYRHEIPGQSLNRAQATALGLADRFELIEDNYAAVNEML 887
 DB 842 RGLYAAFDCTATMKSGNSDVIENEDPGQYTNLHFQAHSMGLSGKFVKFVKAYVEANOML 901
 QY 888 GRPTKVTPTSSKVVGDIALHLVAGVDPAADAPQKYDIPDSVTAFLRGLNGPPEGWPE 947
 DB 902 GDLIKVTPSSKIVGDLAQFVWQNCLSRAEAQAEEISFPSPVVEFLOGYIGIPHGFPPE 961
 QY 948 PLRTRALEG--RSEGGAPLFEVPEEQAHLDAD-----DSKERNRSLNRLFLPKPTEEF 999
 DB 962 PFRSKVLKDLPRIEG-GPGASLPLNLKEKDLIDRHGEVTPEDVLSAAMYPDVFAQF 1020
 QY 1000 LEHRRFGNTSALDDREFFYGLVGEFRLRLPDPVPTLLVRLDAISEPDCKGRNVVAN 1059
 DB 1021 KDTATFGPDLNLRLFLQGPKIATBEFEVELEERKGT-LHIKALAVSDLNRAQRQVFE 1079
 QY 1060 VNGQIRPMRVDRSVESVTATAEKADSSNKGHVAAPAG-VVTVTVAEGDEVKAGDAVAI 1118
 DB 1080 LNCQLRSILVKDQAMKEMHFHPKALKDKVQIGCAPMPGKVIDVKVAAAGAKVKGQPLCV 1139
 QY 1119 IEAMKMEATTITASVDGKIERYVVPAAATKVGEGDLIVVV 1156
 DB 1140 LSAMKMETVVTSPMEGTIRKVVHTKDMTLEGGDLILEI 1177
 RESULT 3
 PYC_HUMAN
 ID PYC_HUMAN STANDARD; PRT; 1178 AA.
 AC P11498; Q16705;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Pyruvate carboxylase, mitochondrial precursor (EC 6.4.1.1) (Pyruvic
 DE carboxylase) (PCB).
 GN PC.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_Taxid=9606;
 RP [1]
 RN SEQUENCE FROM N.A.

CC RX TISSUE=Liver, and Kidney; PubMed=7918683;
 CC RX MEDLINE=95002202; PubMed=7918683;
 CC RA Wang I.D., Du Y., Lisgaris M.V., Mandal S.K., Freytag S.O.,
 CC RA Xing B.-S., Liu T.-C., Kwon M., Patel M.S., Kerr D.S.;
 CC RT "Primary amino acid sequence and structure of human pyruvate
 CC RT carboxylase."; Biochim. Biophys. Acta 1227:46-52(1994).
 CC RL [2]
 CC RN SEQUENCE FROM N.A.
 CC RP TISSUE=Kidney;
 CC RC MEDLINE=94324922; PubMed=8048912;
 CC RA Mackay N., Rigat B., Douglas C., Chen H.S., Robinson B.H.;
 CC RT "cDNA cloning of human kidney pyruvate carboxylase.";
 CC RL Biochem. Biophys. Res. Commun. 202:1009-1014(1994).
 CC RN [3]
 CC RN SEQUENCE FROM N.A.
 CC RP TISSUE=Liver, and Kidney;
 CC RC Walker M.E., Jitrapakdee S., Val D.L., Wallace J.C.;
 CC RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
 CC RN [4]
 CC RP MEDLINE=87212051; PubMed=3555348;
 CC RA Lamhonwah A.-M., Quan F., Gravel R.A.;
 CC RT "Sequence homology around the biotin-binding site of human
 CC RT propionyl-CoA carboxylase and pyruvate carboxylase.";
 CC RL Arch. Biochem. Biophys. 254:631-636(1987).
 CC RN [5]
 CC RN SEQUENCE OF 1135-1178 FROM N.A.
 CC RP MEDLINE=85030380; PubMed=6548474;
 CC RA Freytag S.O., Collier K.J.;
 CC RT "Molecular cloning of a cDNA for human pyruvate carboxylase.
 CC RT Structural relationship to other biotin-containing carboxylases and
 CC RT regulation of mRNA content in differentiating preadipocytes.";
 CC RL J. Biol. Chem. 259:12831-12837(1984).
 CC RN [6]
 CC RN VARIANTS PC DEFICIENCY THR-610 AND ILE-743.
 CC RP MEDLINE=98254451; PubMed=9585612;
 CC RA Carbone M.A., Mackay N., Ling M., Cole D.E.C., Douglas C., Rigat B.,
 CC RA Feigenbaum A., Clarke J.T.R., Haworth J.C., Greenberg C.R.,
 CC RA Seargeant L., Robinson B.H.;
 CC RT "Amerindian pyruvate carboxylase deficiency is associated with two
 CC RT distinct missense mutations."; Am. J. Hum. Genet. 62:1312-1319(1998).
 CC RL Am. J. Hum. Genet. 62:1312-1319(1998).
 CC CC -!- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
 CC INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
 CC ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE
 CC CARBOXYL GROUP TO PYRUVATE IN THE SECOND. CATALYZES IN A TISSUE
 CC SPECIFIC MANNER. THE INITIAL REACTIONS OF GLUCOSE (LIVER, KIDNEY)
 CC AND LIPID (ADIPOSE TISSUE, LIVER, BRAIN) SYNTHESIS FROM PYRUVATE.
 CC -!- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3)(-) = ADP + phosphate +
 CC Oxaloacetate.
 CC -!- COFACTOR: BIOTIN AND MANGANESE.
 CC -!- PATHWAY: GLUCONEOGENESIS AND LIPOGENESIS.
 CC -!- SUBUNIT: HOMOTETRAMER.
 CC -!- SUBCELLULAR LOCATION: Mitochondrial matrix.
 CC -!- DISEASE: DEFICIENCY IN PC CAUSES LACTIC ACIDOSIS, MENTAL
 CC RETARDATION AND DEATH. OCCURS IN THREE FORMS: TYPE A (MILD); TYPE
 CC B (SEVERE NEONATAL) AND A VERY MILD LACTICACIDEMIA.
 CC -!- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOAMIDE TRANSFERASES
 CC AND CARBAMYL PHOSPHATE SYNTHETASES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC EMBL; U04641; AAA99537.1; -;
 CC DR EMBL; S72370; AAA31500.1; -;
 CC DR EMBL; U03891; AAA82937.1; -;
 CC DR EMBL; M26122; AAA36423.1; -;

DR EMBL; K02282; AAA60033.1; -;
 DR PIR; B27883; B27883.
 DR PIR; S01469; S01469.
 DR HSP; P24182; IBNC.
 DR Genew; HGNC:8636; PC.
 DR MIM; 266150; -;
 DR InterPro; IPR001882; Biotin_attach.
 DR InterPro; IPR000089; Biotin_lipoyl.
 DR InterPro; IPR000901; CPSase.
 DR InterPro; IPR000891; HMGL-like.
 DR InterPro; IPR003379; PYC_OADA.
 DR Pfam; PF00289; CPSase_L_chain; 1.
 DR Pfam; PF00364; biotin_lipoyl; 1.
 DR Pfam; PF00682; HMGL-like; 1.
 DR Pfam; PF02436; PYC_OADA; 1.
 DR Pfam; PF02785; Biotin_Carb_C; 1.
 DR Pfam; PF02786; CPSase_LD2; 1.
 DR Pfam; PF02786; CPSase_LD2; 1.
 DR TIGRFAMs; TIGR01235; Pyruv_carbox; 1.
 DR PROSITE; PS00188; BIOTIN; 1.
 KW Ligase; Multifunctional enzyme; Biotin; Manganese; Gluconeogenesis;
 KW ATP-binding; Mitochondrion; Lipid synthesis; Transit peptide;
 KW Disease mutation.
 FT TRANSIT 1 20 MITOCHONDRION (POTENTIAL).
 FT CHAIN 21 1178 PYRUVATE CARBOXYLASE.
 FT DOMAIN 21 549 BIOTIN CARBOXYLASE (BY SIMILARITY).
 FT DOMAIN 550 1000 CARBOXYLTRANSFERASE (BY SIMILARITY).
 FT DOMAIN 1096 1178 BIOTIN CARBOXYL CARRIER PROTEIN
 FT NP_BIND 198 203 ATP (BY SIMILARITY).
 FT ACT_SITE 328 328 BY SIMILARITY.
 FT BINDING 1144 1144 BIOTIN (BY SIMILARITY).
 FT VARIANT 610 610 A -> T (IN PC DEFICIENCY TYPE A).
 FT VARIANT 743 743 M -> I (IN PC DEFICIENCY TYPE A).
 FT VARIANT 743 743 /FTID-VAR_008095.
 FT CONFLICT 225 226 LA -> WP (IN REF. 2).
 FT CONFLICT 352 352 LA -> S (IN REF. 3).
 FT CONFLICT 386 386 EL -> PT (IN REF. 2).
 FT CONFLICT 486 487 EL -> DV (IN REF. 2).
 FT CONFLICT 638 638 P -> R (IN REF. 2).
 FT CONFLICT 729 729 E -> A (IN REF. 2).
 FT CONFLICT 774 775 DT -> AP (IN REF. 2).
 FT SEQUENCE 1178 AA; 129633 MW; 381F527553A20095 CRC64;
 Query Match 43.2%; Score 2531.5; DB 1; Length 1178;
 Best Local Similarity 46.0%; Pred. No. 2.7e-128;
 Matches 541; Conservative 186; Mismatches 415; Indels 35; Gaps 13;
 QY 7 GGLLLKGIITLVSTHTSSTLPA-----FKILVANGETAVRAFRALDTGAA 54
 DB 9 GGLRLGI-----RRTSTAPASPNVRLEYKPIKVMVANGETAIVRFRACTELGR 62
 QY 55 TVAIYPREDRGSPHRSFASAEVRIETGSPVKAYLDIDITGAARKVKADAIYPGYGLS 114
 DB 63 TVAIYSEQDTGMHROKADAYLIGRLAPVQAYLHPDIIVAKENVDVHPGYGLS 122
 QY 115 ENAQLARECAENIGTPTPEVLDLTGDKSRVATAAKKAGLPVL-AESTPSKNIDDIK 173
 DB 123 ERADEAQAQDAGVRFIGPSPVVRKMGDKVEARALAAAGVVPVPGTDAPITSLHEAHE 182
 QY 174 SABGQYPIFVKAVAGGGRGRRFVSSPDELKRLATEAEAEAFADGSGVVERAVINP 233
 DB 183 FSNYTGPIIFKAAVGGGRGRRVHSYELEENYTRAYSEALAFNGALVEFIEKP 242
 QY 234 OHIEVOILGDRGTEVHLVERDCSLORRHQKVVEIAPAOHLDELDRICADAVFCRSI 293
 DB 243 RHIEVOILGQVGNILUHLVERDCSIQRHOKVVEIAPAAHLOPQURTLTSDSVKLAKO 302
 QY 294 GYOGACTVEFLVDEKGNHVFIEPNRIQVETHTVEVDLVKAMRLAAGATKELGL 353
 DB 303 GYENACTVEFLVDRHGKHVFIEVNSRLQVETHTVEVDLVHQAQIHVAEGRSLPDGL 362
 QY 354 TQDKIKTHGAALQCRITTEDPNNNGFRPDTGTITAYRSPGGAGVRLDGAAL-GGEITAHF 412


```
Db 363 ROENIRINGCAIOCRVTTEDPARSFQDPTGRIEVSREGGIRLDNASAFQAVISPHY 422
Qy 413 DMLVKMTCRSGDEETAVARAQALAEFTVSGVATNIGFLRALLREEDFTSKRIATGFIG 472
Db 423 DSSLVKVIANGKHDPHTAKTMSRALAEFRVGRKNTAFQNLNNQOFLAGTVDTQFID 482
Qy 473 DHPHLLQAPPADDOGRILDYLDADVTVNKGHVRP-KDVAAPIDK-LPNIKDLPPLRGRSR 530
Db 483 ENPELFOLRPAQNPAQKLLHLYLGHVMVNGTTPIPYKASPSPTDPVVPVPIGPPAGFR 542
Qy 531 DLKQLGPAAPARLDLQDQDALAVDTTFRDAHOSLLATRVRSFALKPAABEAVAKLTPELL 590
Db 543 DILLREGPEGFARAVRNHPGLLLMDTTFRDAHOSLLATRVTHDLKKIAPVAHNEKSLF 602
Qy 591 SVAWGATVDVAMRFLFDPWRLDELREAMPNVNQLMRLGRNTVGYTPYDSCVCRAF 650
Db 603 SMENMGATFDVAMRFLFYECPRRLQBLRELINIPQMLLRGANAVGYTNDVNVYKFE 662
Qy 651 VKAAASGVDFIRFDALNDVDSQMRPAIDAVLETNTAVAEMAYSGDLSDPNKLYTLD 710
Db 663 CEVAKENGMDVRFVDSLNLPLNMLGMEAGSAG-GVVEAAISYTGVDVADPRTKYSIQ 721
Qy 711 YYLKMAEEIVKSGAHILAIDKMAGLRPAATVTLVTLALREF-DLPVHVHTHTAGQLA 769
Db 722 YYMGLAEELVRAGTHILCIKDMAGLLKPTACTMLVSLSLRDRFPDPLPHIHTDTSAGVA 781
Qy 770 TYFAAAQAGADAVDGSAPLSGTTSPSLSAIIVAFAHTRDTGLSLAEVSDLEPTWEAV 829
Db 782 AMLACAQAGADVVDVADSMGSMGTSPQSMGALVACTRGTPDLDTEVPMERVFDYSEWEGA 841
Qy 830 RGLYLPFESGTPGPTGR--VYRHEIPGQSLNLRQAATLGLADRLPFLIEDNYAAVNEML 887
Db 842 RGLYAFDCTATKSGNSDYENEIPGQVYTNLFQAHSMGLSKGKVEYKKAIVANQML 901
Qy 888 GRPTKVTYPSKVYGDIALHLVAGVDPADPAADQKYDIPDSVIAFLRGLGNPPGQWPE 947
Db 902 GDLIKVTYPSKIVYGDIAQFVQVQNLGSRAAEAQAEELSPRSVVEFLQYIGVPHGQFPE 961
Qy 948 PLRTRAL-----EGRSEKAGPLATEVPEERQAHLD--DQSKERRNSINRLFFPKPTEEFL 1000
Db 962 PFRSKVLKDLPRVEGRGASLPDLQALEKELVDHRGEEVTPEDVLSAAMYPDVFAHEK 1021
Qy 1001 EHRRFRGNTSALDDREFFYGLVREGRETILRLPDVTRPLLVRLDAISEPDDKGMNVVAV 1060
Db 1022 DFTATGFLDSLTRLFLQPKIAEFEVELEGRKT-LHIKALVSDLNRAQORQVFEEL 1080
Qy 1061 NGQIRPMRVNDRSVESVTAEKADSSNKGHVAAFPAG-VVTVVTAEGDEVKAGDAVAII 1119
Db 1081 NGQLRSILVKDQAMKEMHFHPKALDKVKGQIGAPMPGKVIDIKVAVAGAKVAKGQPLCVL 1140
Qy 1120 EAMKMEATITASVDGKTIERVVVVPAATKVGEGDLIVVV 1156
Db 1141 SAMKMETVTSPEMGTIRKVVHTKMDTLEGDDLILEI 1177

RESULT 4
PYCL_YEAST
ID PYCL_YEAST STANDARD; PRT; 1178 AA.
AC P11154;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Pyruvate carboxylase 1 (EC 6.4.1.1) (Pyruvic carboxylase 1) (PCB 1).
GN PYC1 OR PYV OR YGL062w.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=88298805; PubMed=3042770;
RA Lim F., Morris C.P., Occhiodoro F., Wallace J.C.;
```

```
RT "sequence and domain structure of yeast pyruvate carboxylase.";
RL J. Biol. Chem. 263:11493-11497(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX Feuermann M., de Montigny J., Potier S., Souciet J.-L.;
RA "The characterization of two new clusters of duplicated genes
RT suggests a 'Lego' organization of the yeast Saccharomyces cerevisiae
RT chromosomes.";
RL Yeast 13:861-869(1997).
RN [3]
RP SEQUENCE OF 1003-1178 FROM N.A.
RX MEDLINE=87241529; PubMed=3036126;
RA Morris C.P., Lim F., Wallace J.C.;
RY "Yeast pyruvate carboxylase: gene isolation.";
RL Biochem. Biophys. Res. Commun. 145:390-396(1987).
CC -!- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
CC INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
CC ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE
CC CARBOXYL GROUP TO PYRUVATE IN THE SECOND.
CC -!- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3)(-) = ADP + phosphate +
CC oxaloacetate.
CC -!- COFACTOR: BIOTIN AND ZINC.
CC -!- PATHWAY: GLUCONEOGENESIS.
CC -!- SUBUNIT: HOMOTETRAMER.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOAMIDE TRANSFERASES
CC AND CARBAMYL PHOSPHATE SYNTHETASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: J03889; AAA34843.1;
CC EMBL: Z72584; CAA96765.1;
CC PIR: A29233; QYBYP.
CC HSP: P24182; 1BNC.
CC SGD: S0003030; PYC1.
CC InterPro: IPR001882; Biotin_attach.
CC InterPro: IPR000089; Biotin_lipoyl.
CC InterPro: IPR000901; CPSase.
CC InterPro: IPR000891; HMGL-like.
CC InterPro: IPR003379; PYC_OADA.
CC Pfam: PF00289; CPSase_L_chain; 1.
CC Pfam: PF00364; biotin_lipoyl; 1.
CC Pfam: PF00682; HMGL-like; 1.
CC Pfam: PF02436; PYC_OADA; 1.
CC Pfam: PF02785; Biotin_carb_C; 1.
CC Pfam: PF02786; CPSase_L_D2; 1.
CC TIGRfams: TIGR01235; pyruv_carbox; 1.
CC PROSITE: PS00188; BIOTIN; 1.
CC PROSITE: PS00866; CPSASE_1; 1.
CC PROSITE: PS00867; CPSASE_2; 1.
CC Ligase; Multifunctional enzyme; Biotin; Gluconeogenesis; ATP-binding;
KW Zinc; Multigene family.
FT NP_BIND 182 187 ATP (POTENTIAL).
FT ACT_SITE 312 312 BY SIMILARITY.
FT BINDING 1135 1135 BIOTIN (BY SIMILARITY).
FT SIMILAR 160 330 CARBAMOYL PHOSPHATE SYNTHETASES.
FT SIMILAR 350 470 WITH OTHER BIOTIN CARBOXYLASES.
FT SIMILAR 1086 1178 WITH OTHER BIOTIN CARRIER PROTEINS AND
FT WITH LIPOAMIDE ACETYLTRANSFERASE.
FT CONFLICT 462 462 T -> G (IN REF. 1).
FT CONFLICT 493 493 V -> D (IN REF. 1).
FT CONFLICT 595 595 E -> A (IN REF. 1).
FT CONFLICT 619 619 R -> Q (IN REF. 1).
FT CONFLICT 664 664 G -> S (IN REF. 1).
FT CONFLICT 772 772 A -> R (IN REF. 1).
```

[illegible]

DR EMBL; X59890; CAA42544.1; -.
 DR EMBL; Z36087; CAA85182.1; -.
 DR EMBL; U35647; AAC49147.1; -.
 DR PIR; S46094; S46094.
 DR HSP; P24182; IBNC.
 DR SGD; S0000422; PYC2.
 DR InterPro; IPR001882; Biotin_attach.
 DR InterPro; IPR000089; Biotin_lipoyl.
 DR InterPro; IPR000901; CPSase.
 DR InterPro; IPR000891; HMG-L-like.
 DR InterPro; IPR003379; PYC_OADA.
 DR Pfam; PF00289; CPSase_L_chain; 1.
 DR Pfam; PF00364; biotin_lipoyl; 1.
 DR Pfam; PF00682; HMG-L-like; 1.
 DR Pfam; PF02436; PYC_OADA; 1.
 DR Pfam; PF02785; Biotin_carb_C; 1.
 DR Pfam; PF02786; CPSase_L_D2; 1.
 DR TIGRFAMS; TIGR01235; pyruv_carbox; 1.
 DR PROSITE; PS00188; BIOTIN; 1.
 DR PROSITE; PS00866; CPSASE.1; 1.
 DR PROSITE; PS00867; CPSASE.2; 1.
 KW Ligase; Multifunctional enzyme; Biotin; Gluconeogenesis; ATP-binding;
 KW Zinc; Multigene family.
 FT NP_BIND 183 188
 FT ACT_SITE 313 313
 FT BINDING 1136 1136
 FT SIMILAR 161 331
 FT SIMILAR 351 471
 FT SIMILAR 1067 1180
 FT CONFLICT 15 15
 FT CONFLICT 132 132
 FT CONFLICT 238 238
 FT CONFLICT 268 268
 FT CONFLICT 546 546
 FT CONFLICT 642 642
 FT CONFLICT 771 773
 FT CONFLICT 831 831
 FT CONFLICT 839 839
 FT CONFLICT 1001 1001
 FT CONFLICT 1155 1155
 FT CONFLICT 1178 1178
 FT CONFLICT 1180 1180
 FT SEQUENCE 1180 AA; 130166 MW; AD60DA3A60F5E001 CRC64;
 Query Match 41.5%; Score 2436; DB 1; Length 1180;
 Best Local Similarity 45.5%; Pred. No. 3.5e-123;
 Matches 530; Conservative 185; Mismatches 407; Indels 44; Gaps 16;

QY 24 STLPAFKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSPFASAEAVRIGTEG- 82
 DB 15 SLLGEKNKILVANRGEIPIRISAEHLSMRTIAIYSHEDRLSMHRLKADAEYVIGEEQ 74
 QY 83 -SPVAYLDEITGAKKVADAIYPGVGFSENAQLARECAENGITIGPTPEVLDT 141
 DB 75 YTPVGAVLAMEITIEIAKHKKVDFIHPGVGFSENFADKVKAGITWIGPPAEIVDSV 134
 QY 142 GKSRAVTAAKAGLVPVLAEST-PSKNIDDIKVSABGQYPIFVKAVAGGGGRMFVSS 200
 DB 135 GDKVSAHIAAARNVPTVPGTPIETVOEALDFVNEYGYPLIIKAAFGGGGRGMVVRRE 194
 QY 201 PDELARKLATEASREAAFGDGSVYVERAVINPQHTIEVQLIGDRTEGVVHLVERDCSLQR 260
 DB 195 GDOVADAFORATSEARTAFNGTCFVERFLDKPKHIEVOLLDADNHNQVHVLPERDCSVQR 254
 QY 261 RHQKVVEIAPAOHLDPDLDRIDCADAVKFCRSIGYOGAGTVEFLVDEKGNHVFIEINPRI 320
 DB 255 RHQKVVEVAPAKTLPREVRDAILTDVAKLAKVCGYRAGTAEBFLVDNQNRHVFIEINPRI 314
 QY 321 QVEHVTVEEVDVLKAOHRLAAGATLKLGLTQDKIKTHGAALOCRITTEDPNNQFRP 380
 DB 315 QVEHTITEITGIDIVSAQIQIAAGATLQTLGLLQDKITTRGFSIQCRITTEDPSKNFQP 374

QY 381 DTGTTAYRSPGCGAGVRLDGA-AQLGGEITAHFDSMLVKMTCRGSDFETAVARAQALAE 439
 DB 375 DTGRLEVYSAGNGVRLDGGNAYAGATISPHYDSMLVKCSGSIYEVIRRMIRALIE 434
 QY 440 FTVSGVATNIGFLRALLRREEDFTSKRIATGFIQDHPHLLQAPADDEQGLDYLADTVV 499
 DB 435 FRIRGVKTNIPPELLTLLTNPNVTEGTYWTFIDDTQPLQFMVSSNQAKLLHLYLADLAV 494
 QY 500 N---KPHGVKRPDVAAPIDKLPNIKDL-----PLPRGSRDLKQLGPAFAARDL 545
 DB 495 NGSSIKGQIGLPLKLSNP--SVPHLHDAOGNINVTKSPSGMROVLLEKGESEFAK 552
 QY 546 REQDALAVTDTTFRDAHQSLLATRVRSFALKPAEAFAVAKLTPELLSVEAWGATYDAMR 605
 DB 553 ROFNGLTLLMDTTWROAHQSLLATRVTHDLATIAPTTAHALAGAPALECWCWGTFOVAMR 612
 QY 606 FLFEDPWRDLDELREAMPNVNTOMLGRNTVGYTYPDSVCFRAFVKEAASSGVDFIRIF 665
 DB 613 FLHEDPWERLKRSLRSLVPNIPFQMLLRGANGVAYSSLPDNIADHFVKQAKDNGVDIFRVF 672
 QY 666 DALNDVSOMRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYLYLKMAEEIVKSGAH 725
 DB 673 DALNDELUKVGNVAVKAG-GVVEATVCYSGMDLQPGKK-YNLDIYLEVKEIVONGTH 730
 QY 726 ILAIKDMAGLLRPAAVTKLVLTALRREF-DLPVHVHTHTDAGGQATYFAAAQADAVDG 784
 DB 731 ILGINKDMAGTKMPAAKLLIGLSLTRYPDLPIHVHSDSAGTAVASMTACALAGADVVDV 790
 QY 785 ASAPISGTTQSLSLSAIVAFAAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPESGTPGPT 844
 DB 791 AINSMSGLTSQPSINALLASL-EGNIDTGTINVEHVRELDAYNAEMRLLYSCFEADLKGP 849
 QY 845 GRVYRHEIPGGOLSNLRAOATAGLADRFELLEDNYAANEMLRPTKVTPTSSKVVVDLA 904
 DB 850 PEYIOHEIPGGQTLNLLFOAQQLGIGEQAETKRAYREANYLUGDIVKVTPTSKVVVDLA 909
 QY 905 LHLVAGVDPADFAADPQKYDIPDSVIAFLRGLGNPGGWPPLRTRALEGRSEKAPL 964
 DB 910 QFMVSNKLTSDDIRRLANSLLDFPDSVMDFFEGILGQYGFPEPLRSLRNK---RRKL 966
 QY 965 TEVPEEEQHLDAADDKERRNSLNRL-----LFPKPTFEFLHRRRFGNTSAL 1012
 DB 967 TCRPGLELEPFPLE--KIREDLQNREFGIDECDVASNNYPRVYEDFQKIRETYGDSVL 1024
 QY 1013 DDERFPYGLVEGRETLIRLPDVRTPLLRDLDAISEPDDK-GMRNVVANYNGQIRPMVRD 1071
 DB 1025 PTKNFLAPAEPEDEIEVTEQGT-LIILQAVGDLNKKTGQREVYFELNGELRKRIVAD 1083
 QY 1072 RSVESVTATAEKADSSNKGHVAAPFAGV-VTVTVAEGDEVKAGDAVAITEAMKEATITA 1130
 DB 1084 KSONIGSVAKPRADVHDTHQIGAPMAGVITIEVKHKGSLVKKGESIAVLKMKMVMVSS 1143
 QY 1131 SVDGKLERVVVPAATKVEGGDLIVV 1156
 DB 1144 PADGQVKDFIKDGESVDASDLLVL 1169
 RESULT 6
 PYC_PICPA STANDARD; PRT; 1189 AA.
 AC P78992;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Pyruvate carboxylase (EC 6.4.1.1) (Pyruvic carboxylase) (PCB).
 GN PYC1.
 OS Pichia pastoris (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Pichia.
 OX NCBI_Taxid=4922;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98301182; PubMed=9639311;

RA Menendez J., Delgado J., Gancedo C.:
 RT "Isolation of the *Pichia pastoris* pyC1 gene encoding pyruvate
 RT carboxylase and identification of a suppressor of the pyC
 RT phenotype".
 RL Yeast 14:647-654(1998).
 CC -1- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
 CC INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
 CC ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE
 CC CARBOXYL GROUP TO PYRUVATE IN THE SECOND (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3)(-) -> ADP + phosphate +
 CC oxaloacetate.
 CC -1- COFACTOR: BIOTIN AND ZINC.
 CC -1- PATHWAY: GLUCONEOGENESIS.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOAMIDE TRANSFERASES
 CC AND CARBAMYL PHOSPHATE SYNTHETASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Y11106; CAA71993.1; -
 DR HSSP: P24182; LDV1
 DR InterPro: IPR001882; Biotin_attach.
 DR InterPro: IPR000089; Biotin_lipoyl.
 DR InterPro: IPR000901; CPase.
 DR InterPro: IPR000801; HMCL-like.
 DR InterPro: IPR003379; PYC_OADA.
 DR Pfam: PF00289; CPase_L_chain; 1.
 DR Pfam: PF00364; biotin_lipoyl; 1.
 DR Pfam: PF00682; HMCL-like; 1.
 DR Pfam: PF02436; PYC_OADA; 1.
 DR Pfam: PF02785; Biotin_carb_C; 1.
 DR Pfam: PF02786; CPase_L_D2; 1.
 DR TIGRfams: TIGR01235; pyruv_carbox; 1.
 DR PROSITE: PS00188; Biotin; 1.
 DR PROSITE: PS00866; CPase_1; 1.
 DR PROSITE: PS00867; CPase_2; 1.
 DR KW Ligase; Multifunctional enzyme; Biotin; Gluconeogenesis; ATP-binding;
 KW Zinc.
 FT NP_BIND 185 190 ATP (POTENTIAL).
 FT ACT_SITE 315 315 BY SIMILARITY.
 FT BINDING 1140 1140 BIOTIN (BY SIMILARITY).
 SQ SEQUENCE 1189 AA; 131400 MW; 8B6E858079657914 CRC64;

Query Match 40.8%; Score 2393; DB 1; Length 1189;
 Best Local Similarity 45.5%; Pred. No. 7 3e-121;
 Matches 529; Conservative 176; Mismatches 419; Indels 38; Gaps 18;

QY 23 SSTLPAPKLLVANGELIAPRAALETGAATVAIYPREDGSPFRSEASEAVRIGTEC 82
 DB 16 SSLGTHNKKLVANRGEIPRIETAEHLSMNTVAIYSHEDRLSMRLKADEAYVIGRC 75
 QY 83 --SPVKAYLDEIGAAKVKADAIYPGVGLFSENAQLARECAENGITFGTTPVLDL 140
 DB 76 QYSPVQAYLAIDEIKIYAVKHNNMHPGVGFCSENSEFARKVEENGILLWVGSDTVIDA 135
 QY 141 TGKSRVAVTAAGKAGLPVLAESTPSKNIDIVKS---AEGQTYPIFVKAVAGGGRGMRP 197
 DB 136 VGDKVSARNLAYAANVPTV-PGTPGP-IEDVAQATAVEEYGPVVIKAAFGGGGRGMRV 193
 QY 198 VSSPDLKRLKATEASRAEAAGGVSVYERAVINPQHIEVQILGDRGTGEVHLYERDCS 257
 DB 194 VREGDDIEDAFLRASSEAKTAFNGTGFIERFLDKPKHIEVQLLDADNYGNVILHLPERDCS 253
 QY 258 LQRHQRVVEIAPQHLDPRLRACADAVKFCRSTGYCGAGTVEFLVDEKGNHVEIEN 317
 DB 254 VQRHQKVARNCSTAKLPVVEVRNAILDAVKLAKTANYRNAGTAEFLVDSQNRHYFEIN 313

QY 318 PRIQVHTVTBEVTEVDLVKQAMRLAAGATLKELGLTODKIKTHGAALQCRITTEDPNNG 377
 DB 314 PRIQVHTITEITGVDIVAAQIQIAAGASLEQLGLLQEKITTRGFAIOCRITTEDPTKN 373
 QY 378 FRPDGTGTAYRSGGAGVRLDGAALQGE-ITAFHDSMLVKMTCRGSDFFETAVARAQ 436
 DB 374 FQPDGTGKTEVYRSSGGNGVRLDGGNGFAGAVISPHYDSMLVKCSTSGSYEIRRRKMIRA 433
 QY 437 LAETVSVGAVNIGFLRALLREEDFTSKRIATGFTGDHPHLLOAPPADDEOGRILDLAD 496
 DB 434 LVEFRIRGVKINIFELLALLTHPVMTSECTWTTFIDDTPELFKILTSQNRQAQLLAYLGD 493
 QY 497 VTVN-----KPHGVPRKQVAAIDKLPNIKD--LPLPRGSRDRLKQLQPAAPAF 543
 DB 494 LAVNGSSIKGOIGLUPKLH--READIPSITDINGVDIVISIPPDGWRQFLLEKPEQFAQ 551
 QY 544 DLREODALAVTDTTFRDAHQSLLATRVRSFAKPAEAAVAKLTPELLSVSEAWGATYDVA 603
 DB 552 QYRAPFPGIMINDTWRDAHQSLLATRVTHDLLNIATPATSYALHHAFALCEMGGATEDVS 611
 QY 604 MRFLEDPWDRDELREAMPNNVNIQMLLRGNTVGYTPYDPSVCRAFVYKEAASSGVDIR 663
 DB 612 MRFLEDPWDRDELREAMPNNVNIQMLLRGNTVGYTPYDPSVCRAFVYKEAASSGVDIR 671
 QY 664 IFDALNDVSOIRPAIDAVLENTNVAEAMAYSGDLSDPNEKLYTLDYYLKMAREIVKSG 723
 DB 672 VFDALNDIEQLKVGVDVAVKKAG-GVVEATMCSGDMKPKKK-YNLEYIINLATEIVEMG 729
 QY 724 AHILAIKDMAGLLRPAAYTKLVLTALRREF-DLPVHVHTHTDAGGOLATYFAAAGADAV 782
 DB 730 THILAVKDMAGTLKPTAAKQLISALRRKFPSLPINVHTDSAGTGVASWVACARAGADV 789
 QY 783 DGASAPLSGTTSPSLSAIVAAFAHTRDRDTGLSLEANSVDELEPYEAVRGLYLPESGTPG 842
 DB 790 TVRVNMSGMTSOPSMSAFIASL-DGEIETGIPENAREIDAYWAEMLLLSCFEADLKG 848
 QY 843 PFGVRYRHEIPGGQSLNLRQAATALGLADRFELIBDNVAANVMGLRPTKVTTPSKVVD 902
 DB 849 PDPEVYQHEIPGGQSLNLRQAATALGLADRFELIBDNVAANVMGLRPTKVTTPSKVVD 908
 QY 903 LALHLVGAGVDPADPAADPOKYDIPDSVIAFLRGLNPPGGMPELPTRALEGSRSE--- 959
 DB 909 LAQFVMSNKLSESDVERLASELDFPDSVLDFFEGLMGTPYGGFPEPLRTNVISGRKRLT 968
 QY 960 GKAPLTPVEEEQA---HLDADDSKERNNSL-NRLLFPKPTFEFLHRRRFGNTSALDDR 1015
 DB 969 SRPGLTLEPYNIPALREDLEAFRSKVTENDVASYNMYPKYVEAYKKQQLXGDLVLPTR 1028
 QY 1016 EPFY--GLVEGRETLIRLPDVRTPLLRDLDAISE--PDDKGMNVNVANVNGQILRPMVRDR 1072
 DB 1029 NFLSPPKIDEERHVTIVTETRTKTLIIKMAEGELSSQSGTREVYFELNGEMKRVTVEDK 1088
 QY 1073 SVESVTAETAEKADSSNKGHVAFAGVVT-VTVASGDEVKAGDAVAITEAMKMEATITAS 1131
 DB 1089 NGAVETITRPKADAHNPNEIGAPMAGVVVEVVRHENGVEKKGDPIAVLSAMKMEVISP 1148
 QY 1132 VDGKTIERVVPAATKVEGGDLI 1153
 DB 1149 VAGRIQIATKENDSVSDASDLI 1170
 RESULT 7
 PYCA_METJA STANDARD: PRT; 501 AA.
 ID PYCA_METJA AC Q58626;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Pyruvate carboxylase subunit A (EC 6.4.1.1) (Pyruvic carboxylase A).
 OS PYCA OR M1229.
 GN Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 CC Methanocaldococcaceae; Methanocaldococcus.

D	b		2	FNVKLIANGETAIRILIRCBWELGTLKVAVYSEADKKSLHATLDAEAYCIG-PAPRAKSY	60
Q	y		89	LDIDEITIGAARKVADAIYPGYFLSENAQLARECAENGITFIPTPEVDLTQDKSR	148
		I::I::	I::I::	I::I::I::I::I::I::I::	I::I::I::I::I::I::I::
D	b		61	LNIADILNVAERAKYDAIHPGYGFLAENAFARAVKVKAGFEFIGPNDPAIEAMGS	120
Q	y		149	TAAKKAGLPVLAEPTSPS-KNIDDDIVKSAGQOTPIFYKAVAGGGGRMRFPSS	207
		I::I::I::I::	I::I::I::I::	I::I::I::I::I::I::I::	I::I::I::I::I::I::I::
D	b		121	KIMKKAGVPLIPGSEGATEIDEAETIAEATGFFPVVVKASAGGGGGMGSVAYSKE	180
Q	y		208	ATASREAEAAAGDGSVVVERAVINPOHEVQILGDRTEGVVHLRYEDCSLORRHOK	267
		I::I::I::I::I::I::I::	I::I::I::I::I::I::I::	I::I::I::I::I::I::I::	I::I::I::I::I::I::I::
D	b		181	IESARNIAKSAFGDPVTFTYEKLNPRIETIQILGDKHGNIIHLGDEKCSIQRRHOK	240
Q	y		268	IAPAQHLDPELRDICADAVKFCRSIGYGAGTVEFLVDEKGNHVFTFEMPRIQVEHT	327
		I::I::	I::I::	I::I::I::I::I::I::I::	I::I::I::I::I::I::I::
D	b		241	EAPSPTWTELREMERCEAAIKAGKAINYSAGTVEFLY-ENGNYFYLEMNTRIQVEHT	299
Q	y		328	EEVTEVDLVKAQMRLAAGATLKELGLTDQTKTHGAALQCRIITTEDPNNGFRPD	387
		I::I::I::I::I::I::I::	I::I::I::I::I::I::I::	I::I::I::I::I::I::I::	I::I::I::I::I::I::I::
D	b		300	EQVTVGDLVKAMIKTAAG---BELTKQEDVKIRGHATEACINAEDPLNDFVPCPK	356
Q	y		388	YRSPGAGVRLDGAAOLGCGETTAHFDSMLVKWTCRGSDFTAVARAQALAEFTVS	447
		I::I::I::I::I::I::I::	I::I::I::I::I::I::I::	I::I::I::I::I::I::I::	I::I::I::I::I::I::I::
D	b		357	YRSPGGPVRIODSGVYGAEIPPYDSMIKTLYTNSREBAIRMKRALLREVVI	416
Q	y		448	NIGFLRALIREEDFTSKRIATFGIDGHPHLQ	479
		I::I::I::I::I::I::I::	I::I::I::I::I::I::I::	I::I::I::I::I::I::I::	I::I::I::I::I::I::I::
D	b		417	NIPFHRAVLEENFLKGNISTHYVQNMIKLR	448
<hr/>					
RESULT 8					
ID	ACC_ANASP	STANDARD;	PRT;	447 AA.	
AC	O06862:				
DT	01-FEB-1995 (Rel. 31, Created)				
DT	01-FEB-1995 (Rel. 31, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	Biotin carboxylase (EC 6.3.4.14) (A subunit of acetyl-CoA				
DE	carboxylase (EC 6.4.1.2)) (ACC).				
GN	ACC OR ALR0939,				
OS	Anabaena sp. (strain PCC 7120).				
OC	Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.				
OX	NCBI_TaxId:103690;				
RN	[1]				
RN	SEQUENCE FROM N.A.				
RX	MEDLINE=93352435; PubMed=8102363;				
RA	Gornicki P., Scappino L.A., Haselkorn R.;				
RA	"Genes for two subunits of acetyl coenzyme A carboxylase of Anabaena				
RT	sp. strain PCC 7120: biotin carboxylase and biotin carboxyl carrier				
RT	protein.";				
RL	J. Bacteriol. 175:5268-5272(1993).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=21595285; Pubmed=11759840;				
RA	Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,				
RA	Matanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,				
RA	Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,				
RA	Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,				
RA	Yasuda M., Tabata S.;				
RT	"Complete genomic sequence of the filamentous nitrogen-fixing				
RT	Cyanobacterium Anabaena sp. strain PCC 7120."				
RL	DNA Res. 8:205-213(2001).				
CC	-I- FUNCTION: THIS PROTEIN IS A COMPONENT OF THE ACETYL COENZYM				
CC	CARBOXYLASE COMPLEX; FIRST, BIOTIN CARBOXYLASE CATALYZES THE				
CC	CARBOXYLATION OF THE CARRIER PROTEIN AND THEN THE TRANS-CARBOXYLASE				
CC	TRANSFERS THE CARBOXYL GROUP TO FORM MALONYL-CoA.				
CC	-I- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + Co(2)				
CC	= ADP + phosphate + carboxybiotin-carboxyl-carrier protein.				
CC	-I- PATHWAY: Long-chain fatty acid biosynthesis; first step.				
CC	-I- SUBUNIT: ACETYL-CoA CARBOXYLASE IS AN HETEROHEXAMER OF BIOTIN				
CC	CARBOXYL CARRIER PROTEIN, BIOTIN CARBOXYLASE AND THE TWO SUBUNITS				
CC	OF CARBOXYL TRANSFERASE IN A 2:2 COMPLEX.				

CC -1- SIMILARITY: TO OTHER BIOTIN-DEPENDENT ENZYMES AND CARBAMOYL-
CC PHOSPHATE SYNTHETASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL: L14862; AAB51770.1; -;
DR EMBL: AP003584; BAB72896.1; -;
DR HSSP: P24182; 1BNC.
DR InterPro: IPR004549; AccC.
DR InterPro: IPR000901; CPSase.
DR Pfam: PF02785; CPSase_L_chain; 1.
DR Pfam: PF02785; Biotin_carb_C; 1.
DR Pfam: PF02786; CPSase_L_D2; 1.
DR TIGRFAMS: TIGR00514; accC; 1.
DR PROSITE: PS00866; CPSASE.1; 1.
DR PROSITE: PS00867; CPSASE.2; 1.
DR Fatty acid biosynthesis; Ligase; Biotin; ATP-binding;
KW Complete proteome. 168 ATP (BY SIMILARITY).
FT NP_BIND 163 293 BY SIMILARITY.
FT ACT_SITE 293 293
SQ SEQUENCE 447 AA; 49104 MW; 8A541B38B39E00F9 CRC64;

Query Match 17.6%; Score 1032.5; DB 1; Length 447;
Best Local Similarity 48.4%; Pred. No. 1.8e-48;
Matches 216; Conservative 71; Mismatches 148; Indels 13; Gaps 4;

QY 29 FKKILVANGETAVRAAFRAALTEGATVAIYPRDGRSGFHRSPASAVRIGTGGSPVKAY 88
DB 3 FDKILIANRGEIAVRLIACEEMGIAIVSHVTDNRALHVLADEAVCIQ-SPASAKSY 61
QY 89 LDIDEIIGAARKVADAIYPGYGLSNAQARECAENGITFIGPTPEVLDLTKSRVAV 148
DB 62 LNIPIIATAALTRNSAIIHGYGLSNAKFAEACDHHIAFTGPTPEAIRLMGDKSTAK 121
QY 149 TAARKAGLPVLAES-----TPSKNIDIVKSAEQTYPIFKAVAGGGGRMFVSPDE 203
DB 122 ETMOKAGVPTVPGSEGLVTEQEGLE---LAKDIGYVPMIKATAGGGGRMLVRSPE 177
QY 204 LRKLATEASREAAAFDGGVYVERAVINPQHIEVQILGDRTEGVHLYERDCSLQRRHQ 263
DB 178 FVKFLAAQGEAGAGAGNAGVYIEKFTERPRIHFQILADNYGNV IHLGERDCSIQRRNQ 237
QY 264 KVEIAPQAHLDPRLDRICADAVKFKRSIGYCGAGTVEELVDEKGNHVFIEKNPRIOVE 323
DB 238 KLEEAPSPALDSDLRKMGQAQVAAKAAQFINYTCAGTIEFLDRSGQFYEMENNTRIOVE 297
QY 324 HTVTEEVTEVDLVKAQMLAAGATLKLGLTQDKIKTKHGAALOCRIITPDNNGFRPDGT 383
DB 298 HPVTMTVGVLLVEQIRIAGERLR--LTQDQVVLRLHGAIECRINAEDPDHDFRAPG 354
QY 384 TITAYRSPGGAGVLDGAQALGGEITAHFDSMLVKMTCRGSDETAVARAORALAEFTVS 443
DB 355 RISGLPPGGVGRVDSHVYDQIIPYDLSLIGKLVWGPDRATARNMKRALRECAIT 414
QY 444 GVATNIGFLRALLREEDFTSKRIATGFI 471
DB 415 GLPTTIGFHQIMENPQFLOGNVSTSFV 442

RESULT 9
PYCA_ARCFU STANDARD; PRT; 506 AA.
AC O30019;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pyruvate carboxylase subunit A (EC 6.4.1.1) (Pyruvic carboxylase A).

GN PYCA OR AF0220.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_taxid=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.J., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.A., Graham M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham M.D., Sutton G.G., Gill S.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Adams M.D., Lofus B.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Badger J.H., Glodek A., Zhou L.,
RA Peterson S., Reich C.I., McNeil L.K., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
RA Sadov P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
CC -1- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
CC INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
CC ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE CARBOXYL
CC GROUP TO PYRUVATE IN THE SECOND.
CC -1- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3)(-) = ADP + phosphate +
CC Oxaloacetate.
CC -1- COFACTOR: ATP, MAGNESIUM (OR MANGANESE OR COBALT), PYRUVATE AND
CC BICARBONATE (BY SIMILARITY).
CC -1- PATHWAY: GLUCONEOGENESIS.
CC -1- SUBUNIT: HETEROCTAMER OF FOUR A AND FOUR B SUBUNITS (BY
CC SIMILARITY).
CC -1- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOAMIDE TRANSFERASES
CC AND CARBAMYL PHOSPHATE SYNTHETASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation;
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE001090; AAB91012.1; -;
CC HSSP: P24182; 1BNC.
CC TIGR: AF0220; -;
DR InterPro: IPR004549; AccC.
DR InterPro: IPR000901; CPSase.
DR Pfam: PF02789; CPSase_L_chain; 1.
DR Pfam: PF02785; Biotin_carb_C; 1.
DR Pfam: PF02786; CPSase_L_D2; 1.
DR TIGRFAMS: TIGR00514; accC; 1.
DR PROSITE: PS00866; CPSASE.1; 1.
DR PROSITE: PS00867; CPSASE.2; FALSE_NEG.
DR Ligase; Multifunctional enzyme; Gluconeogenesis; Magnesium; Pyruvate;
KW ATP-binding; Complete proteome.
FT NP_BIND 161 290 ATP (POTENTIAL).
FT ACT_SITE 290 290 POTENTIAL.
SQ SEQUENCE 506 AA; 57438 MW; 39B765F319235AD1 CRC64;

Query Match 17.4%; Score 1020; DB 1; Length 506;
Best Local Similarity 47.9%; Pred. No. 1e-47;
Matches 212; Conservative 76; Mismatches 149; Indels 6; Gaps 4;

QY 29 FKKILVANGETAVRAAFRAALTEGATVAIYPRDGRSGFHRSPASAVRIGTGGSPVKAY 88
DB 2 FSKILVANGETAVRVMRACRELIGTKTVGVYSSADKRAFHRVYADECYIG-KADPRDSY 60
QY 89 LDIDEIIGAARKVADAIYPGYGLSNAQARECAENGITFIGPTPEVLDLTKSRVAV 148
DB 61 LNIIDRIEIVAKKSGAEIHPGYGLFALNAEFAECRCEEGIVFIGPSPEVIRIAGSKVRSR 120

```
QY 149 TAAKAGLPVLAETSTKSNIDDIKVAEGQTYPIFKAVAGGGGMRGVSSPDELKLA 208
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 ESMQAGVPVPCSKPIDTDEAKENAKIGYPVAVKASGGGGIGIVVNSQEELEAF 180
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 209 TEASREAEAAFGSGVYERAVINPOHIEVQIIGDRTGTVVHLYERDCSLQRHOKVWEI 268
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 RSKKLGESYFKDSTVYLEKYLARPHIEVQIILADQHGNIHLGERECSTQRHOKLIEE 240
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 269 APAQHLDPELRDRIKADVKFCRSIGYQAGTVFELVDEKGNHVFTEMPIQVEHTVTE 328
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 241 APSPALNEEMREELGKLAVKAGREIGTNGTFFELY-ENGNYFFLEINSLRQVEHTITE 299
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 329 EYTEVDLVKAQMLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNNGFRPDTGITAY 388
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 300 VVTGIDIVKQIRIAYG---EELRHGOEDVAIRGHAIECRINAEDPVN-FYPRSGRILHY 355
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 389 RSPGGAGVRLDGAQALGGETITAFHDSMLVWKTCRGSDFETAVARAQALAEFTVSGVATN 448
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 356 RSPGGGIRVDSGIHMGRIPEYDSMISKLIAIYGTREAEIARMKRALYIEYIEGVTN 415
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 449 IGLRALLREEDFTSKRIATGFI 471
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 416 IPHFVAVLNDSEFVRGNIHTKFV 438
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
PYCA.METH
ID PYCA.METH STANDARD; PRT; 491 AA.
AC 027939;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE Pyruvate carboxylase subunit A (EC 6.4.1.1) (Pyruvic carboxylase A).
GN PYCA OR MTH1917.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Delta H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Alldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiواني N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
RN [2]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-17.
RC STRAIN=Delta H;
RX MEDLINE=98148063; PubMed=9478969;
RA Mukhopadhyay B., Stoddard S.F., Wolfe R.S.;
RT "Purification, regulation, and molecular and biochemical
RT characterization of pyruvate carboxylase from Methanobacterium
RT thermoautotrophicum strain deltaH.";
RL J. Biol. Chem. 273:5155-5166(1998).
CC -!- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
CC INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
CC ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE CARBOXYL
CC GROUP TO PYRUVATE IN THE SECOND. THE MAXIMUM ACTIVITY IS AT PH 8
CC AND 60 DEGREES CELSIUS.
CC -!- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3)(-) = ADP + phosphate +
CC oxaloacetate.
CC -!- COFACTOR: ATP, MAGNESIUM (OR MANGANESE OR COBALT), PYRUVATE AND
CC BICARBONATE.
CC -!- ENZYME REGULATION: INHIBITED BY ADP AND ALPHA-KETOGLUTARATE.
CC -!- PATHWAY: GLUCONEOGENESIS.
```

DE carboxylase (EC 6.4.1.2)) (ACC).
GN ACC OR HI0972
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OX Haemophilus.
NX NCBI_TaxID=727;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN-Rd / KW20 / ATCC 51907;
RC MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kienle A., Sutton G., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKelvey K., Bult C.J., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.";
RL Science 269:496-512(1995).
CC -1- FUNCTION: THIS PROTEIN IS A COMPONENT OF THE ACETYL COENZYME A CARBOXYLASE COMPLEX; FIRST, BIOTIN CARBOXYLASE CATALYZES THE CARBOXYLATION OF THE CARRIER PROTEIN AND THEN THE TRANSCARBOXYLASE TRANSFERS THE CARBOXYL GROUP TO FORM MALONYL-COA (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2) -> ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
CC -1- PATHWAY: Long-chain fatty acid biosynthesis; first step.
CC -1- SUBUNIT: ACETYL-COA CARBOXYLASE IS AN HETERODIMER OF BIOTIN CARBOXYL CARRIER PROTEIN, BIOTIN CARBOXYLASE AND THE TWO SUBUNITS OF CARBOXYL TRANSFERASE IN A 2:2 COMPLEX (BY SIMILARITY).
CC -1- SIMILARITY: TO OTHER BIOTIN-DEPENDENT ENZYMES AND CARBAWOYL-PHOSPHATE SYNTHETASES.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U32778; AAC22632.1; -;
DR HSSP; P24182; 1BNC.
DR TIGR; HI0972; -;
DR InterPro; IPR004549; AccC.
DR Pfam; PF00289; CPSase_Lchain; 1.
DR Pfam; PF02785; Biotin_carb_C; 1.
DR Pfam; PF02786; CPSase_LD2; 1.
DR TIGRFAMs; TIGR00514; accC; 1.
DR PROSITE; PS00866; CPSASE_1; 1.
DR PROSITE; PS00867; CPSASE_2; 1.
DR Fatty acid biosynthesis; Ligase; Biotin; ATP-binding;
KW Complete proteome.
FT NP_BIND 163 168 ATP (POTENTIAL).
FT ACT_SITE 292 292 POTENTIAL.
SQ SEQUENCE 448 AA; 49108 MW; 28497E2A31ED96D1 CRC64;
Query Match 16.2%; Score 950; DB 1; Length 448;
Best Local Similarity 48.3%; Pred. No. 4.9e-44;
Matches 212; Conservative 57; Mismatches 152; Indels 18; Gaps 9;
QY 30 KKLIVANGETAVRAAETGAATVAIYPRDRGSRFSRFASEAVRIGTEGSPKAYL 89
DB 3 EKXVIANRGEIATRLIRACKELGKIKTVAHSTADRLDKHLLADETICIGPAPS-AKSYL 61
QY 90 DIDEITIGAKKVKADAIYPGVGFLENALARECAENGTITFGTPEVLDLTGDKSRAVT 149
DB 62 NIPATIAAEVGTGDAIHPGVGFLENADFAEQVRSRGTFTIGTADVIRLMGDKVSAIK 121
QY 150 AKKAGLPVLAEST-PSKNIDDIIVS---AEGQTYPIEVKAVAGGGGRGMRVSPDELRL 205

DB 122 AMKAGVPCVPGSGVPSN--DIANKKETAKRIGYPIIKASGGGGRGMRVVRSEDALE 179
QY 206 KLATEASREAAFGDGVYVERAVINPOHIEVILGDTGEVHLXERDCSLORRHOKV 265
DB 180 ESIAMTKAEAKAANNDWMYERLENPRHVEIQVLADTHGNVYLAERDCSMQRHOKV 239
QY 266 VEIAPQHLDPEDRLRI---CADAVKFCRSICYGAGTVEFLVDEKGNHVFLEMPIQOV 322
DB 240 VBEATAPGTEVRRDIDGSRCA---CVEIGYRGAGTFFLY-ENGFEYFIEMMTRIQV 295
QY 323 EHTVTEEVTEVLDVKAOMRLAAGATLKEGLTQDKIKTHGAALQCRITTEDPNNFRPDT 382
DB 296 EHPVTEMITGVLDVKEQLRIAGL---PTSFQEDIKVKGHAMECRINAEDPKT-FLPSP 351
QY 383 GTITAYRSRPGAGVRLDGAQGLGTEITAHFDSMLVKMTCRGSDFFETAVARAQALAEFTV 442
DB 352 GKVNHLSRPSGLGVWRWDVSHVGYTPVPHYDSMIAKLITYGDTREVAIRRMQNALSETII 411
QY 443 SGVATNIGELFRALLREEDF 461
DB 412 DGIKTNIPHLHELLEDENF 430
RESULT 12
ACCC_ECOLI
ID ACCC_ECOLI STANDARD; PRT; 449 AA.
AC P24182.
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 43, Last annotation update)
DE Biotin carboxylase (EC 6.3.4.14) (A subunit of acetyl-CoA carboxylase (EC 6.4.1.2)) (ACC).
DE ACCC OR PABG OR B3256.
GN Escherichia coli.
OS Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
[1]
RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RP STRAIN-K12;
RC MEDLINE=92052166; PubMed=1682920;
RX Kondo H., Shiratsuchi K., Yoshimoto T., Masuda T., Kitazono A.,
RA Tsuru D., Anai M., Sekiguchi M., Tanabe T.;
RT "Acetyl-CoA carboxylase from Escherichia coli: gene organization and nucleotide sequence of the biotin carboxylase subunit.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:9730-9733(1991).
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE=92112819; PubMed=1370469;
RA Li S.-J., Cronan J.E. Jr.;
RT "The gene encoding the biotin carboxylase subunit of Escherichia coli acetyl-CoA carboxylase.";
RL J. Biol. Chem. 267:855-863(1992).
[3]
RN SEQUENCE FROM N.A.
RP Best E.A., Knauf V.C.;
RA Submitted (XXX-1992) to the EMBL/GenBank/DBJ databases.
[4]
RN SEQUENCE FROM N.A.
RX STRAIN-K12 / MGI655;
RC MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-VIDES J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
[5]
RN SEQUENCE OF 1-12.
RP STRAIN=K12 / EMG2;
RC MEDLINE=97443975; PubMed=9298646;
RX Link A.J., Robison K., Church G.M.;

RT "Comparing the predicted and observed properties of proteins encoded
 RL in the genome of *Escherichia coli* K-12.";
 RT Electrophoresis 18:1259-1313(1997).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
 RX MEDLINE=94347758; PubMed=7915138;
 RA Waldrop G.L., Rayment I., Holden H.M.;
 RT "Three-dimensional structure of the biotin carboxylase subunit of
 RL acetyl-CoA carboxylase.";
 RL Biochemistry 33:10249-10256(1994).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
 RX MEDLINE=20283656; PubMed=10821865;
 RA Thoden J.B., Blanchard C.Z., Holden H.M., Waldrop G.L.;
 RT "Movement of the biotin carboxylase B-domain as a result of ATP
 binding.";
 RL J. Biol. Chem. 275:16183-16190(2000).
 CC -!- FUNCTION: THIS PROTEIN IS A COMPONENT OF THE ACETYL COENZYME A
 CC CARBOXYLASE COMPLEX; FIRST, BIOTIN CARBOXYLASE CATALYZES THE
 CC CARBOXYLATION OF THE CARRIER PROTEIN AND THEN THE TRANSCARBOXYLASE
 CC TRANSFERS THE CARBOXYL GROUP TO FORM MALONYL-COA.
 CC -!- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
 CC = ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
 CC -!- PATHWAY: Long-chain fatty acid biosynthesis; first step.
 CC -!- SUBUNIT: ACETYL-COA CARBOXYLASE IS AN HETEROHEXAMER OF BIOTIN
 CC CARBOXYL CARRIER PROTEIN, BIOTIN CARBOXYLASE AND THE TWO SUBUNITS
 CC OF CARBOXYL TRANSFERASE IN A 2:2 COMPLEX.
 CC -!- SIMILARITY: TO OTHER BIOTIN-DEPENDENT ENZYMES AND CARBAMOYL-
 CC PHOSPHATE SYNTHETASES.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; M79446; AAA23748.1; -
 DR EMBL; M80458; AAA23409.1; -
 DR EMBL; M83198; AAA23746.1; -
 DR EMBL; U18997; AAA58059.1; -
 DR EMBL; AE000404; AAC76288.1; -
 DR PIR; JS0632; JS0632.
 DR PIR; JS0687; JS0687.
 DR PDB; 1BNC; 30-AUG-95.
 DR PDB; 1DV1; 09-JUN-00.
 DR PDB; 1DV2; 09-JUN-00.
 DR Ecogene; EG10276; accC.
 DR InterPro; IPR004549; AccC.
 DR InterPro; IPR000901; CPSase.
 DR Pfam; PF00289; CPSase_L_chain; 1.
 DR Pfam; PF02785; Biotin_carb_C; 1.
 DR Pfam; PF02786; CPSase_L_D2; 1.
 DR TIGRFAMS; TIGR00514; accC; 1.
 DR PROSITE; PS00866; CPSASE.1; 1.
 DR PROSITE; PS00867; CPSASE.2; 1.
 KW Fatty acid biosynthesis; Ligase; Biotin; ATP-binding; 3D-structure;
 KW Complete proteome.
 FT NP_BIND 163 168 ATP (POTENTIAL).
 FT ACT_SITE 292 292 POTENTIAL.
 FT CONFLICT 260 261 CA -> SR (IN REF. 2).
 FT CONFLICT 313 313 L -> M (IN REF. 1).
 SQ SEQUENCE 449 AA; 49320 MW; 68C55F10ACB4F170 CRC64;

Query Match 16.0%; Score 940; DB 1; Length 449;
 Best Local Similarity 46.8%; Pred. NO. 1.7e-43;
 Matches 203; Conservative 66; Mismatches 155; Indels 10; Gaps 6;

QY 31 KILVNRGCIARAFRAALETCAATVAIYPREDRCFSFRHSFASAVRIGTCGSPVKAYLD 90
 DB 4 KIVIANRGCIARILRACKELGIKTVAHVSSADRLKHLVLLADETVCIGPAPS-VKSYLN 62

QY 91 IDEIIAAKKVADAIYPGYFLSENAQLARECAENGITFIGPTPEVLDLTGDKSAVTA 150
 DB 63 IPATISAAEITGVAIHPCYGFSENAFAEQVERSGFIFIGPKAETIRLMGDKVSAIAA 122
 QY 151 AKKAGPLVLAESTPSKNIDDIKVS---AEGQTYPIFVAVAGGGGRGMRFPVSSPDELK 207
 DB 123 MKKAGVPCVPGSDPLG-DMDKNRAIKRIGYPVITKASGGGGGRGMRVVRGDAELAQ 181
 QY 208 ATEASREAAAFCDGSGVYVERAVINPOHIEVQILGDRTEGVVHLYERDCSLORRHQV 267
 DB 182 ISMTRAEAKAAFSNDVMYMEKYLENPRHVEIQVLADQGNIAIYLAERDCSMQRHOKV 241
 QY 268 IAPAQHLDELDRICADAVKFCRSIGYOGAGTVEELVDEKGNHVIENMPRIQVHTVT 327
 DB 242 EAPAGITFELRYIGERCACAKVDIGYRGAGTFELF-ENGFEYFLENTRIQLQVEHPVT 300
 QY 328 EEVTEVDLVKAQMRLAAGATLKGTLGTQDKIKTHGAALOCRTTPEDPNNNGFRPDTGT 387
 DB 301 EMITGVDLILKEQLRIAG--QPLSIKQBEVHVGRHAVECRINAEDPNT-FLPSPGKITR 356
 QY 388 YRSPGCGAYRLDGAALGGEITAHFDSMLVKTCTGSDPETAFAVARAQLAEFTVSGVAT 447
 DB 357 FHAPGGFVGRWESHYAYVTVPYDYMIGKLCICYGENRDVAIARMKNALQELIIDGIKT 416
 QY 448 NIGFLKALLREEDF 461
 DB 417 NVDLQIRINDENF 430

RESULT 13

ACCC_ECO57
 ID ACCC_ECO57 STANDARD; PRT; 449 AA.
 AC O8X9B6;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE Biotin carboxylase (EC 6.3.4.14) (A subunit of acetyl-CoA
 DE carboxylase (EC 6.4.1.2)) (ACC).
 GN ACCC OR 24616 OR ECS4128.
 OS *Escherichia coli* O157:H7.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC *Escherichia*.
 OX NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Postfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7.";
 RL Nature 409:529-533(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
 RT "Complete genome sequence of enterohaemorrhagic *Escherichia coli*
 RT O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:11-22(2001).
 CC -!- FUNCTION: THIS PROTEIN IS A COMPONENT OF THE ACETYL COENZYME A
 CC CARBOXYLASE COMPLEX; FIRST, BIOTIN CARBOXYLASE CATALYZES THE
 CC CARBOXYLATION OF THE CARRIER PROTEIN AND THEN THE TRANSCARBOXYLASE
 CC TRANSFERS THE CARBOXYL GROUP TO FORM MALONYL-COA.
 CC -!- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
 CC = ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
 CC -!- PATHWAY: Long-chain fatty acid biosynthesis; first step.

QY 30 KKILVANGEIAVRAALETGAATVAIYPREDGSRFHRSEAFASAVRIGTEGSPVKAYL 89
 Db 3 EKVLIANRGEIALRLRKLKELGKIKTVAHVSTADRELMLHLSLADESVICG-PAPATQSYL 61
 QY 90 DIDEIIGAIAKKVADAIYPGVGFLSENAQLARECAENGTIFGPTPEVLDLTGDKSRVAVT 149
 Db 62 QIPALIAAAEVGTAAIHPGVGFLAENADFAEQIERSFTFVGPAAEIRLMDGDKVSADK 121
 QY 150 AAKKAGLPVL--AESTPSKNTDDIVKSAEGOTYFVKAAGGGGRGMRFFVSSPDELRLK 207
 Db 122 AMKRAGVTPVSGDPLDEDEETALAIAREVGYPIIIKAAGGGGRGMRVYVDESELK 181
 QY 208 ATEASREAEAFGDSVYVERAVINPQHIEVOILGDRTGVEVHVLYERDCSLQRHKKVVE 267
 Db 182 AKLTRTEAGAAFGNPMVYLEKFLTPRHVEVQVLSGOGNATHLGRDCSLQRHKKVIE 241
 QY 268 IAPAOHLPELRDRICADAVFCRSIGYGAGTVEFLVDEKGNHVFIEWNPRIQVEHVT 327
 Db 242 EAPAPGIDEKARQEVFARCVQACIEIGYRGAGTPEFLY-ENGRFFYFIEMNTRVQVHEPVS 300
 QY 328 EEVTEVDLVKQMRLLAAGATLKLGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITA 387
 Db 301 EWVTGDIVKEMLRATSG---EKLSIQEDVVIRGHALECRINAEDPKT-TWMPSPGKVKH 356
 QY 388 YRSPGAGVRLDGAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQALAEFTVSGVAT 447
 Db 357 FHAPGNGVRVDSHLYSGVSPNPVDSLVGVITVYAGDRDEALARMNALDELIVDGIKT 416
 QY 448 NIGFLRALLREDF 461
 Db 417 NTELHKDLVRDAAF 430

RESULT 15
 MCCA_HUMAN
 ID MCCA_HUMAN STANDARD; PRT; 725 AA.
 AC Q9RQ3; Q9H59; Q9NS97;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Methylcrotonyl-CoA carboxylase alpha chain, mitochondrial precursor
 DE (EC 6.4.1.4) (3-Methylcrotonyl-CoA carboxylase 1) (MCCase alpha
 DE subunit) (3-Methylcrotonyl-CoA:carbon dioxide ligase alpha subunit).
 GN MCC1 OR MCCA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP PUBMED=11170888;
 RX Gallardo M.E., Desviat L.R., Rodriguez J.M., Esparza-Gordillo J.,
 RA Perez-Cerdá C., Perez B., Rodriguez-Pombo P., Criado O., Sanz R.,
 RA Morton D.H., Gibson K.M., Le T.P., Ribes A., Rodriguez de Cordoba S.,
 RA Ugarte M., Penalba M.A.;
 RT "The molecular basis of 3-methylcrotonylglycinuria, a disorder of
 RT leucine catabolism".
 RL Am. J. Hum. Genet. 68:334-346(2001).
 RN [2]
 RP SEQUENCE FROM N.A., AND VARIANTS MCGI ARG-325 AND SER-385.
 RX MEDLINE=21295033; PubMed=11401427;
 RA Obata K., Fukuda T., Morishita R., Abe S., Asakawa S., Yamaguchi S.,
 RA Yoshino M., Ihara K., Murayama K., Shigemoto K., Shimizu N., Kondo I.;
 RT "Human biotin-containing subunit of 3-methylcrotonyl-CoA carboxylase
 RT gene (MCCA): cDNA sequence, genomic organization, localization to
 RT chromosomal band 3q27, and expression".
 RL Genomics 72:145-152(2001).
 RN [3]
 RP SEQUENCE FROM N.A., AND VARIANT MCGI PHE-535.
 RX MEDLINE=21299419; PubMed=11406611;
 RA Holzinger A., Roeschinger W., Lagler F., Mayerhofer P.U., Lichtner P.,
 RA Kattenfeld T., Thuy L.P., Nyhan W.L., Koch H.G., Muntau A.C.,
 RA Roscher A.A.;

"Cloning of the human MCCA and MCB genes and mutations therein reveal
 the molecular cause of 3-methylcrotonyl-CoA: carboxylase
 deficiency".
 Hum. Mol. Genet. 10:1299-1306(2001).
 [4]
 RP SEQUENCE FROM N.A., AND VARIANTS MCGI VAL-289; SER-385; PRO-437 AND
 RP HIS-532.
 RX PubMed=11181649;
 RA Baumgartner M.R., Almashanu S., Suormala T., Obie C., Cole R.N.,
 RA Packman S., Baumgartner E.R., Valle D.;
 RT "The molecular basis of human 3-methylcrotonyl-CoA carboxylase
 RT deficiency".
 J. Clin. Invest. 107:495-504(2001).
 [5]
 RP SEQUENCE FROM N.A.
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzûki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
 RA Matsumura Y., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
 RA Nakamura Y., Nagahara K., Masuho Y., Sasaki N.;
 RT "NEDO human cDNA sequencing project".
 RN Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 [6]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skeletal muscle;
 RA Strausberg R.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ATP + 3-methylcrotonyl-CoA + HCO(3)(-) - ADP +
 CC phosphate + 3-methylglutaconyl-CoA.
 CC -1- COFACTOR: Biotin.
 CC -1- PATHWAY: Leucine catabolism.
 CC -1- SUBUNIT: Probably a dodecamer composed of six biotin-containing
 CC alpha subunits and six beta subunits.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
 CC -1- DISEASE: Defects in MCC1 are the cause of 3-
 CC methylcrotonylglycinuria type I (MCGI, CGA or CG2). MCGI is a
 CC recessive disease that is characterized by muscular hypotonia and
 CC atrophy, probably of spinal origin.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL; AF310972; AAG53095.1; -;
 EMBL; AB029826; BAA99407.1; -;
 EMBL; AF297332; AAK67986.1; -;
 EMBL; AF310339; AAG50245.1; -;
 EMBL; AK023051; BAB14377.1; -;
 EMBL; BC004214; AAH04214.1; -;
 EMBL; BC004187; AAH04187.1; -;
 EMBL; HGNC:6936; MCC1.
 MIM; 210200; -;
 DR InterPro; IPR001882; Biotin_attach.
 DR InterPro; IPR000089; Biotin_lipoyl.
 DR InterPro; IPR000901; CPSase.
 DR Pfam; PF00289; CPSase_L_chain; 1.
 DR Pfam; PF00364; biotin_lipoyl; 1.
 DR Pfam; PF02785; biotin_carb_C; 1.
 DR Pfam; PF02786; CPSase_L_D2; 1.
 DR PROSITE; PS00188; BIOTIN; 1.
 DR PROSITE; PS00867; CPSASE_2; 1.
 KW Mitochondrion; Ligase; Biotin; ATP-binding; Transit peptide;
 KW Disease mutation; Polymorphism.
 FT TRANSIT 1 47
 FT CHAIN 48 725 MITOCHONDRION (POTENTIAL).
 FT NP_BIND 209 214 METHYLCROTONYL-CoA CARBOXYLASE ALPHA
 FT ACT_SITE 339 339 CHAIN.
 FT BINDING 681 681 ATP (POTENTIAL).
 FT DOMAIN 538 541 BIOTIN (BY SIMILARITY).
 FT POLY-SER. 541

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2003, 18:57:30 ; Search time 110 seconds
(without alignments)
2167.241 Million cell updates/sec

Title: US-09-974-973-2
Perfect score: 5865
Sequence: 1 MTATLGLGLLKGIITLST.....RVVVPATKVEGGDLIVVVS 1157

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5759	98.2	1140	2	O54587
2	5297.5	90.3	1139	2	O8RQL2
3	3797	64.7	1124	16	Q9RK64
4	3786.5	64.6	1127	2	Q9F843
5	3713.5	63.3	1127	16	P95127
6	3484	59.4	1124	2	Q50450
7	2535.5	43.2	1150	16	Q9K9M0
8	2530	43.1	1148	16	Q9KWD4
9	2523.5	43.0	1180	13	Q9DDT1
10	2509	42.8	1147	2	P94448
11	2507.5	42.8	1192	3	Q9HES8
12	2494	42.5	1146	16	Q8Y846
13	2491	42.5	1146	16	Q92CW1
14	2490.5	42.5	1144	16	Q97FR7
15	2489	42.4	1150	16	Q99UY8
16	2485.5	42.4	1193	3	O93918

17	2474.5	42.2	1174	16	Q8UBX3
18	2473.5	42.2	1152	16	Q98F27
19	2465.5	42.0	1158	16	Q8YJ20
20	2461	42.0	1175	5	O17732
21	2447	41.7	1185	3	Q90UE1
22	2441	41.6	1181	5	Q9XZ00
23	2439	41.6	1185	3	P78822
24	2435.5	41.5	1175	3	Q8X1T3
25	2435	41.5	1152	16	O92L13
26	2433.5	41.5	1196	5	Q9I7E9
27	2417.5	41.2	1195	5	O16921
28	2413	41.1	1137	16	Q9CHQ7
29	2405	41.0	1137	2	Q9RAT6
30	2401	40.9	1154	2	Q9Y740
31	2119.5	36.1	984	2	Q9XBJ1
32	2021	34.5	935	11	O62043
33	1933.5	33.0	920	2	Q9KWU5
34	1071.5	18.3	477	16	O67449
35	1040	17.7	472	16	O67483
36	1005.5	17.1	453	2	O54755
37	982.5	16.8	384	2	O07640
38	979.5	16.7	471	16	Q9HTD0
39	978.5	16.7	448	16	Q55160
40	973.5	16.6	1095	16	Q9I3U4
41	970.5	16.5	1338	5	Q9GQ08
42	969.5	16.5	493	17	O8TSX0
43	966.5	16.5	493	1	Q9HH17
44	966	16.5	573	2	Q9FAF5
45	963.5	16.4	1078	16	Q9A3J0

ALIGNMENTS

RESULT 1

O54587	ID	O54587	PRELIMINARY;	PRT;	1140 AA.
AC	O54587;				
DT	O1-JUN-1998	(Tremblrel. 06, Created)			
DT	O1-JUN-1998	(Tremblrel. 06, Last sequence update)			
DT	O1-JUN-2002	(Tremblrel. 21, Last annotation update)			
DE	Pyruvate carboxylase.				
GN	pyc.				
OS	Corynebacterium glutamicum (Brevibacterium flavum).				
OC	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;				
OC	Actinomycetales; Corynebacterineae; Corynebacteriaceae;				
OC	Corynebacterium.				
OX	NCBI_TaxID=1718;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=21253;				
RA	Koffas M.A.G.; Ramamoorthi R.; Pine W.A.; Sinskey A.J.;				
RA	Stephanopoulos G.;				
RL	Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=ATCC 13032;				
RA	Peters-Wendisch P.G.; Kreutzer C.; Kalinowski J.; Patek M.; Sahm H.;				
RA	Eikmanns B.J.;				
RL	Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.				
CC	-!- COFACTOR: BIOTIN (BY SIMILARITY).				
DR	EMBL; AF038548; AAB92588.1;				
DR	EMBL; Y09548; CAA70739.1;				
DR	HSSP; P24182; IBNC.				
DR	InterPro; IPR001882; Biotin_attach.				
DR	InterPro; IPR000089; Biotin_lipoyl.				
DR	InterPro; IPR000901; CPsase.				
DR	InterPro; IPR000891; HMGL-like.				
DR	InterPro; IPR003379; PYC_OADA.				
DR	Pfam; PF02785; Biotin_carb_C; 1.				
DR	Pfam; PF00364; biotin_lipoyl; 1.				
DR	Pfam; PF00289; CPsase_L_chain; 1.				
DR	Pfam; PF02786; CPsase_L_D2; 1.				


```
Db 423 FNVGVATNIGFALLREEDFTKRRIDGTGSHQHLLQAPADDEQGRILEYLAADVTV 482
QY 500 NKPHGVPRKDVAAIDKLNKIDPLPRGSRDLKOLGPAAPARDLREODALAVDTTFR 559
Db 483 NKPHGERP-ETARPIEKLFEVENIPLPGSRDLKOLGPEGFARDLREODALAVDTTFR 541
QY 560 DAHQSLATRVRSFALKPAEAAVAKLTPELLSVGAAGATYDVAMRFLFEDPWDRLELR 619
Db 542 DAHQSLATRVRSFALTPAARAVAKLTPELLSVGAAGATYDVAMRFLFEDPWDRLELR 601
QY 620 EAMPNNIOMLGRNTVGYTYPDSVCRAFVKEAAASSGVDIFRIPDALNDVSOQMPAID 679
Db 602 EAMPNNIOMLGRNTVGYTYPDSVCRAFVQEAASGVDIFRIPDALNDVSOQMPAID 661
QY 680 AVLENTVAEAMAYSGDLSNPNEKLYTLDYVYKMAEIVKSGAHLAKOMAGLLRPA 739
Db 662 AVLEGTCTVAEAMAYSGDLSNPGEKLYTLDYVYLNLAQIVDSGAHLAKOMAGLLRRA 721
QY 740 AVTKLVTLRRFDPVHVHTHTAGGQATYFAAAQAGADAVDGAASPLSGTTSQPSLS 799
Db 722 AAPKLVTLRRFDPVHVHTHTAGGQATYLAANAGADAVDAASAPLSGTTSPSMS 781
QY 800 AIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGILYLPFESGTPGPTGRVYRHEIPGQOLS 859
Db 782 AIVAAFAHTRRDTGLNQLQAVSDLEPYWEAVRGILYLPFESGTPGPTGRVYRHEIPGQOLS 841
QY 860 LRAQATAGLADRFEIENYAAVNMELGRPTKPTSSKVVGDALHLVYAGVDPADFAA 919
Db 842 LRAQAVAGLADRFEIENYAAVNMELGRPTKPTSSKVVGDALHLVYAGVSPEDFAA 901
QY 920 DPKYDIPDSVIAFLRGELGNPPGWPPELRTALRGESGKAPLTPVEPEEQAHLDAD 979
Db 902 DPKYDIPDSVIAFLRGELGTPPGWPPELRTALRGESGKAPLTPVEPEEQAHLDSD 961
QY 980 SKERNSLNRLFLFPKPTFEFLHRRFRFGNTSALDREFFYGLVEGRETLIRLPDVTPL 1039
Db 962 SAERGTNLNRLFLFPKPTFEFLHRRFRFGNTSALDREFFYGLKEGREELIRLTGVSTPMV 1021
QY 1040 VRLDAISEDDKGMNVANVANGQIRPMVRDRSVESVTATKADSSNKGHVAAPFAGV 1099
Db 1022 VRLDAVSEDDKGMNVVNVANGQIRPKVRDRSVESVTATKADATNKGHVAAPFAGV 1081
QY 1100 VTTVAEGDEVKAGDAVIAIEAMKMEATITASVDGKIERYVVPAAFKVEGGDLIVVVS 1157
Db 1082 VTTVAEGDEIKAGDAVIAIEAMKMEATITAPVDGVIDRVVVPAAFKVEGGDLIVVVS 1139

RESULT 3
Q9RK64
ID Q9RK64 PRELIMINARY; PRT: 1124 AA.
AC Q9RK64;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Pyruvate carboxylase.
GN SC00546 OR ScF11.26C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycetaceae; Streptomycetaceae; Streptomyces.
OX NCBI_taxid=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Seeger K.J., Harris D.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
```

```
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Wieser H.M., Denapaita D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL: AL132662; CAB59603.1; -.
DR HSSP: P24182; IBNC.
DR InterPro: IPR000089; Biotin_lipoyl.
DR InterPro: IPR000901; CPSase.
DR InterPro: IPR000891; HMGL-like.
DR InterPro: IPR003379; PYC_OADA.
DR Pfam: PF02785; Biotin_carb_C.1.
DR Pfam: PF00364; biotin_lipoyl.1.
DR Pfam: PF00289; CPSase_L_D2.1.
DR Pfam: PF02786; CPSase_L_D2.1.
DR Pfam: PF00682; HMGL-like.1.
DR Pfam: PF02436; PYC_OADA.1.
DR TIGRFAMs: TIGR01235; pyruv_carbox.1.
DR PROSITE: PS00867; CPSASE_2; UNKNOWN_1.
SQ SEQUENCE 1124 AA; 121071 MW; 1FEF4C4643954D31 CRC64;
```

```
Query Match 64.7%; Score 3797; DB 16; Length 1124;
Best Local Similarity 65.2%; Pred. No 6e-195;
Matches 735; Conservative 156; Mismatches 227; Indels 10; Gaps 6;
QY 29 FKKILVANGEIAVRAFALEATGAATVAIYPREDRGSHRSPASEAVRITGEGSPVKAY 88
Db 2 FRKVLNVRGEIAIRAFRAGYELGARTVAVFPHEDRNSLRKADAEYIEGEGHPVAY 61
QY 89 LDIDELIIGAKKADAIYPGYGLSENALARECAENGITFGPTPEVLDTLGDGSRAY 148
Db 62 LSVEEIVRAARRAGADAVYPGYGLSENPELARACEAEGITFVGPSARILELTGNKARAY 121
QY 149 TAAKAGLPVLAESTPSKNIDDIKSAEGQTYPIFKAVAGGGGRMRFVSSPDELRLKLA 208
Db 122 AAAREAGVPVLGSSAPSTDVDELVRAADVDGFPFVKAVAGGGGRMRRVPEPAQLREAI 181
QY 209 TEASREAAAFGSGVYVRAVINPQIHIEVQILGDRGTGEVHLYERDCSLQRHKKVVEI 268
Db 182 EAASREAAAFGSDTVLEKAVVEPRHIEVQILADGEGDVIHLFERDCSVQRHKKVIEL 241
QY 269 APAQHLDPELRDRICADAVKFCRSICYGAGTVEFLVDEKGNHVFTEMNPRIOVEHTVTE 328
Db 242 APAPNLDPALRERICADAVNFARQIGYRNAGTVEFLVDRDGNHVFTEMNPRIOVEHTVTE 301
QY 329 EVTEVDLVKQMLAAGATLKEGLTQDKTKTHGAALQCRITTEDPNNGFRPDGTGTITAY 388
Db 302 EYTDVDLVQSOLRIAGQTLADLGLAQENITLRGAALQCRITTEDPANGFRPDGTGQISAY 361
QY 389 RSPGGAGVRLD-GAALGGEITAHFDSMLVKMTCRSDFETAVARAQALAEFTVSGVAT 447
Db 362 RSPGGGIRLDGGTTHAGTEISAHFDMSMLVKLCRGRDFTTAVNARRAVERIRGVAT 421
QY 448 NIGFLRALLREEDFTSKRIATATGIGDHPHLLQAPPADDEQGRILDYLAADVTKNPKHCVRP 507
Db 422 NIPFLQAVLDDPDPFQAGRVTTSFIEQRPHLLTARHSADRGTKLLITYLAADVTKNPKHGERP 481
```

```
Qy 508 KDVAAPIDKLPNIKDLPLPGSRDRLLKQLGPAFAFARDLRQDQDALAVTDTTFRDAHQSLLA 567
Db 482 -ELVDPLTKLPTASAGEPAGSROLLAELGEPFARLRRESSTIGVYDTTFRDAHQSLLA 540
Qy 568 TRVRSFALKPAAEAVAKLTPELLSVEAWGATYDVAMRFLFEDPWDRDLREAMPNVNI 627
Db 541 TRVRTKMLAVPVVARTLPOLLSECGGATYDVALLFLAEDPWERLAALREAVPNLCL 600
Qy 628 QMLLRGRNTGYTTPYDPSVCRFAFYKEAASSGVDFRIFDALNDVDSQMRPAIDAVLETNTA 687
Db 601 QMLLRGRNTGYTTPYTEVTDADFQEAATGIDIFRIFDALNDVEQMRPAIEAVROTGSA 660
Qy 688 VAEVAMAYSGDLSDPNEKLYTLDYLLKMAEIVKSGAHILAIKDMAGLLRPAATVLTVA 747
Db 661 VAEVAMAYSGDLSDPNEKLYTLDYLLKMAEIVKSGAHILAIKDMAGLLRPAATVLTVA 720
Qy 748 LRREFDLPVHVHTDTHAGGOLATYFAAAQAGADAVDCASAPLSGTTSSQPSLSAIVAAFDH 807
Db 721 LRREFDLPVHVHTDTHAGGOLATYFAAAQAGADAVDCASAPLSGTTSSQPSLSAIVAAFDH 780
Qy 808 TRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTPGRVYRHEIPCGOLSNLRACATL 867
Db 781 TERPTGLDQAVGDLPEYWEAVRGLYLPFESGTPGTPGRVYRHEIPCGOLSNLRACATL 840
Qy 868 GLADRFELIEDNYAANVMGLRPTKVPSSKVYGDIALHVLGAGVDPADPAADPKYDIP 927
Db 841 GLGRFEDIENYAAANVMGLRPTKVPSSKVYGDIALHVLGAGVDPADPAADPKYDIP 900
Qy 928 DSVTAFLRGLNPGGWPPELPTRALGRSEKAPLTVPEVEEQAHLDADDSKERNLSL 987
Db 901 DSVVGLRGLNPGGWPPELPTRALGRSEKAPLTVPEVEEQAHLDADDSKERNLSL 955
Qy 988 NRLFPKTEFELEHRRFGNTSALDREFFYGLVGEHRETLRL-PDVRTPPLVRLDAIS 1046
Db 956 NRLFPKTEFELEHRRFGNTSALDREFFYGLVGEHRETLRL-PDVRTPPLVRLDAIS 1013
Qy 1047 EPDDKMRNVANVNGQIRPMVRDRSVESVTATAEKADSSNKHGHVAAFPAGVYTVTVAE 1106
Db 1014 DADSRGMRNVANVNGQIRPMVRDRSVESVTATAEKADSSNKHGHVAAFPAGVYTVTVAE 1073
Qy 1107 GDEYKAGDAVAITAEKMEATITASVDGKIERYVVPVPAATKVEGGDLIV 1154
Db 1074 GDEYKAGDAVAITAEKMEATITASVDGKIERYVVPVPAATKVEGGDLIV 1121

RESULT 4
Q9F843 PRELIMINARY; PRT; 1127 AA.
AC Q9F843
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DE 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE Pyruvate carboxylase (EC 6.4.1.1).
GN PYC.
OS Mycobacterium smegmatis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_FaxID=1772;
RN [J].
RP SEQUENCE FROM N.A.
RX MEDLINE=20374587; PubMed=10913817;
RA Mukhopadhyay B., Purwantinil E.;
RT "Pyruvate carboxylase from Mycobacterium smegmatis: stabilization,
RT rapid purification, molecular and biochemical characterization and
RT regulation of the cellular level."
RL Biochim. Biophys. Acta 1475:191-206(2000).
DR EMBL; AF262949; AAG30411.1; -.
DR HSSP; P24182; 1BNC.
DR InterPro; IPR001882; Biotin_attach.
DR InterPro; IPR000089; Biotin_lipoyl.
DR InterPro; IPR000091; CPase.
DR InterPro; IPR000891; HMGL-like.
DR InterPro; IPR003379; PYC_OADA.
```

```
DR InterPro; IPR000634; S/T_dehydrtse.
DR Pfam; PF02785; Biotin_carb_C; 1.
DR Pfam; PF00364; biotin_lipoyl; 1.
DR Pfam; PF00289; CPase_L_chain; 1.
DR Pfam; PF02786; CPase_L_D2; 1.
DR Pfam; PF00682; HMGL-like; 1.
DR Pfam; PF02436; PYC_OADA; 1.
DR TIGRFAMS; TIGR01235; pyruv_carbox; 1.
DR PROSITE; PS00188; BIOTIN; UNKNOWN.1.
DR PROSITE; PS00867; CPASE_2; UNKNOWN.1.
DR PROSITE; PS00165; DEHYDRATASE_SER_THR; UNKNOWN.1.
KW Ligase.
SQ SEQUENCE 1127 AA; 120765 MW; 18C132C48425C67B CRC64;

Query Match 64.6%; Score 3786.5; DB 2; Length 1127;
Best Local Similarity 66.2%; Pred. No. 2.2e-194;
Matches 747; Conservative 142; Mismatches 233; Indels 7; Gaps 6;

Qy 31 KILVANRGEIIVRAFPAALETGAATVAIYPREDGSFHRSPFASAVRIGCTESPVKAYLD 90
Db 4 KVLVANRGEIIVRAFPAALETGAATVAIYPREDGSFHRSPFASAVRIGCTESPVKAYLD 63
Qy 91 IDEITGAKKVADAIYPGYGLFSENAQLARECAENGTFFIGPTPEVLDTGDKSRVTA 150
Db 64 VDEIRVAKHGADAVYPGYGLFSENPDLAAKCAEAGITFFVGPSEVLDTGDKSRVTA 123
Qy 151 AKKAGLPVLAESTPSKNIDIVKSAGQYPIFVXAVAGGGGRGMRFFVSSPDELRLATE 210
Db 124 ARAAGLPVLSSEPSVDELMAAAADMEFFLVKAVSGGGGRGMRFFVSSPDELRLATE 183
Qy 211 ASREAAFGDGSVYVERAVINPOHIEVOILGDRTGVEVHLYERDCSLORRHQKVEIAP 270
Db 184 ASREAAFGDGSVYVERAVINPOHIEVOILGDRTGVEVHLYERDCSLORRHQKVEIAP 243
Qy 271 AOHLPDELDRICADAKFCRSIGYGAGTVEFLVDEKGNHVFIECNPRIOVEHTVTEV 330
Db 244 APNLSDELQQICADAVAFARQIGYSCAGTVEFLVDEKGNHVFIECNPRIOVEHTVTEV 303
Qy 331 TEVDVLKQAMRLAAGATKELGTDQKIKTHGAALQCRTTDPNNGRPDGTITAVRS 390
Db 304 TDVDLVSSOLRIAAGETLADLGSQDLRVVGAAMQCRITTEVPANGFRPDGTITAVRS 363
Qy 391 PGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDFAVARAQRALAEFTVSGVATNIG 450
Db 364 PGAGVRLDGGTNLGAEISAHFDSMLVKMTCRGSDFAVARAQRALAEFTVSGVATNIP 423
Qy 451 FLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLDVTVNKPGRPKDV 510
Db 424 FLOAVIDDPDFRAGRVTTSFIDRPHLLTSRSPADRGTRILNYLADITVKNKPHGERSPTV 483
Qy 511 AAPIDKLPNIKDL--PLPRGSRDLKQLGPAFAFARDLRQDQDALAVTDTTFRDAHQSLLAT 568
Db 484 -YPQRLPPL-DLQAPPAGSKRLVELGPGFAGWLRESKAVGVTDTTFRDAHQSLLAT 541
Qy 569 RYRSFALKPAAEAVAKLTPELLSVEAWGATYDVAMRFLFEDPWDRDLREAMPNVNIQ 628
Db 542 RYRTTGLMLVAPYVARSMFOLLSEICWGATYDVALLFLKEDPWERLAALRESVFNICLQ 601
Qy 629 MLLRGRNTGYTTPYDPSVCRFAFYKEAASSGVDFRIFDALNDVDSQMRPAIDAVLETNTAV 688
Db 602 MLLRGRNTGYTTPYDPSVCRFAFYKEAASSGVDFRIFDALNDVDSQMRPAIDAVLETNTAV 661
Qy 689 AEWAMAYSGDLSDPNEKLYTLDYLLKMAEIVKSGAHILAIKDMAGLLRPAATVLTVA 748
Db 662 AEWAMAYSGDLSDPNEKLYTLDYLLKMAEIVKSGAHILAIKDMAGLLRPAATVLTVA 721
Qy 749 RREFDLPVHVHTDTHAGGOLATYFAAAQAGADAVDCASAPLSGTTSSQPSLSAIVAAFDH 808
Db 722 RREFDLPVHVHTDTHAGGOLATYFAAAQAGADAVDCASAPLSGTTSSQPSLSAIVAAFDH 781
Qy 809 RRDGTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTPGRVYRHEIPCGOLSNLRACATL 868
Db 782 QYDTCGLDLRAVCDLEPYWEAVRGLYLPFESGTPGTPGRVYRHEIPCGOLSNLRACATL 841
```



```

Db 838 IALGLGRFEDIEAYAGADRVGLRVKVTPTSKVVGDLALALVAGVSADFAADPARF 897
QY 925 DIPDSVIAFLRGELNGPWPPELRTALREGRSEKAPLTPVEPEEQAHLDAADSKERR 984
Db 898 GIPESVLGFLRGELGDPGWPPELRTAALAGRAAR-PTAQAADDEIALSSVGAK-RQ 955
QY 985 NSLNRLLFPKPTPEFLHRRFRNGTSALDDREFFYGLVEGRETLIRLPDVRTPLLVRLDA 1044
Db 956 ATLNRLLFPSPKTEFNEHREAYAGDTSOLSAQFFYGLRQGEHVRKL-ERGVLLGLEA 1014
QY 1045 ISEPDGMRNVANVNGQLRPMRDRSRESVETATAEKADSNKNGHVAAPFAGVTVTV 1104
Db 1015 ISEPDGMRVTMCLNGQLRPLVLRDRSISASVPAEAKADRGPNGHIAAPFAGVTVGV 1074
QY 1105 AEGDEVKAGDAVATIEAMKWEATITASVCGKIERVVVPAATKVEGDLIVVVS 1157
Db 1075 CVGERVAGGOTIATIEAMKWEAPITAPVAGTVVERVAUSDQAEGGDLVVVS 1127

RESULT 6
ID Q50450 PRELIMINARY; PRT; 1124 AA.
AC Q50450;
DF 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Pyc.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OX Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
RN NCBI_TaxID=1773;
RN [1]
RA SEQUENCE FROM N.A.
RA Smith D.R.;
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RA Robison K.;
RA SEQUENCE FROM N.A.
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
RL EMBL; U00024; AAA0948.1; -.
DR HSP; P24182; IBNC.
DR InterPro; IPR001892; Biotin_attach.
DR InterPro; IPR000089; Biotin_lipoyl.
DR InterPro; IPR000901; Cpsase.
DR InterPro; IPR000891; HMGL-like.
DR InterPro; IPR003379; Pyc-OADA.
DR InterPro; IPR000634; S/T_dehydratse.
DR Pfam; PF02785; Biotin_carb_C; 1.
DR Pfam; PF00364; biotin_lipoyl; 1.
DR Pfam; PF00289; Cpsase_L; 1.
DR Pfam; PF02786; Cpsase_L-D2; 1.
DR Pfam; PF00682; HMGL-like; 1.
DR Pfam; PF02436; Pyc-OADA; 1.
DR TIGRFAMs; TIGR01235; pyruv_carbox; 1.
DR PROSITE; PS00188; BIOTIN; UNKNOWN.1.
DR PROSITE; PS00867; CPSASE.2; UNKNOWN.1.
DR PROSITE; PS00165; DEHYDRATASE_SER_THR; 1.
SQ SEQUENCE 1124 AA; 120672 MW; 07E1692E12203491 CRC64;

Query Match 59.4%; Score 3484; DB 2; Length 1124;
Best Local Similarity 62.2%; Pred. No. 3 5e-178;
Matches 698; Conservative 151; Mismatches 251; Indels 22; Gaps 9;

QY 29 FKKILVANRGEIAVRAFALETCATVATYIPREDGRSPHRSFASAVRIGTEGSPVXAY 88
Db 2 FSKVLVANRGEIAVRAFALETCATVATYIPREDGRSPHRSFASAVRIGTEGSPVXAY 61
QY 89 LDIDEIIGAKKVADAIYGYGFLSENAQLARECAENGTITFGPTPEVLDTGDKSRV 148
Db 62 LSVDEIVATARRAGADAIYGYGFLSENPDLAAACAAAGISFVGPSEAFLVLEAGNSRAI 121
QY 149 TAAKAGLPVLAESTPSKNIDIVKSAEGQTYPIFKAVAGGGGGRMRFVSSPDELKLA 208

```

```

Db 122 AAAREAGLPVLMSSAPSASVDELLSVAAGMPPELFVKAVAGGGGRMRVYDIAALPEAI 181
QY 209 TEASREAAAFGDSGVYVERAVINPQHIEVQILGORTGEVHLYERDCSLQRRHKQVVEI 268
Db 182 EAASREAESAFGDPVTYLEQAVINPRHIEVQILADNLGDIVHLYERDCSVQRHKQVIEL 241
QY 269 APAQHLDPRLRORICADAVKFCRSIGYOCAGVVEFLVDEKGNHVFIEIENPRIOVHTYTE 328
Db 242 APAPHLDIAELRYKCMKDAVAFARHIGYSCAGTEFLIDERGEYFIEIENPRIOVHTYTE 301
QY 329 EYTEDVLAQOMRLAAGATLKLGLTQDKIKTHGAALOCRITTEDPNNGFRDPTGTITAY 388
Db 302 EITDVLVASQIRIAGELEQGLRQEDIAFHGAALOCRITTEDPANGFRP-TRAGSAR 360
QY 389 RSPGAGVRLDCAAOQGGEITAFHFSMLVKM-----TCRGSOFETAVARAQALAEFTVS 443
Db 361 COPPAVPVSAMTAAP-----TWRRNOPVLRLHAGQADLSGRDLPYAVSRARIAEFRI 415
QY 444 GVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLOAPPADDEOGRILDYLAADVTVNKP 503
Db 416 GVSTNIPFLQAVLDPPDFRAGRVTTSFIDERPOLLTARASADRGTKIINFLADVTVN 475
QY 504 GVRPKDVAAPIDKLNKIDPL-----PRGSRDLKQLGPAFAFARDLREQDALAVDTTFR 559
Db 476 GSRPSTI-YPDKLP---DLDLRAAPPAGSKQRLVKLGPGEFARWLRESAAVGVDTTFR 531
QY 560 DAHQSLLATRVRSFALKPAEAAVAKLTPELSSVEAWGGATYDVAMRFLFEDPDWRLDEL 619
Db 532 DAHQSLLATRVRTSGLSRVAPYLARTMPOLLSVECGGATYDVALRFLKEDPWERLATLR 591
QY 620 EAMPNVNIQMLRGRNTVGYTPYDPSVCFRAVKEAASSGVDFIRFDALNDVSONRPAID 679
Db 592 AAMPNICLQMLLRGNTVGYTPYPEIVTSFVQEATATGIDIFRFDALNNIESNRPAID 651
QY 680 AVLENTVAEAVAMAYSGDSDPNKLYTLDYLLKMAEIVKSGAHILAIDKMAGLLRPA 739
Db 652 AVRETSARAEVAMCYTGLDTPDGEQLTYLKLQIVDGAHVLAKIDMAGLLRPP 711
QY 740 AVTKLVTLALREFDLPVHVITHDTAGGOLATYFAAAGADAGADVAGASAPLSTTSQPSLS 799
Db 712 AAQRLVSALRSRFDLPVHLTHDTGPGQLASYAAAHAGADAGVAGAAAPLACTTSQPAL 771
QY 800 AIVAAFAHTRDTGLSLEAVSDLEPYWEAVRGLYLPFSGTPGTPGRVYRHEIPGGQLSN 859
Db 772 SIVAAAHTEYDTGLSLSAVCALEPYWEALRKVYAPFESGLPGTPGRVYHHEIPGGQLSN 831
QY 860 LRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPTSSKVVGDLALHLVAGVDPAFAA 919
Db 832 LRQQAIALGLGRFEEIEEAYAGADRVGLRVKVTPTSKVVGDLALALVAGVSADFAAS 891
QY 920 DPQKYDIPDSVIAFLRGELNGPWPPELRTALREGRSEKAPLTPVEPEEQAHLDAAD 979
Db 892 DPARTGIPESVLGFLRGELGDPGWPPELRTAALAGRAAR-PTAQAADDEIALSSVG 950
QY 980 SKERNLSNRLLPKPTPEFLHRRFRNGTSALDDREFFYGLVEGRETLIRLPDVRTPL 1039
Db 951 AK-RQATLNRLLFPSPKTEFNEHREAYAGDTSOLSAQFFYGLRQGEHVRKL-ERGV 1008
QY 1040 VRLDAISEPDGMRNVANVNGQLRPMRDRSRESVETATAEKADSNKNGHVAAPFAGV 1099
Db 1009 IGLEAISEPDGMRVTMCLNGQLRPLVLRDRSISASVPAEAKADRGPNGHIAAPFAGV 1068
QY 1100 VTVTVAGDEVKAGDAVATIEAMKWEATITASVCGKIERVVV 1141
Db 1069 VTVGVCGVERVAGGOTIATIEAMKWEAPITAPVAGTVVERVAV 1110

RESULT 7
QYK9M0
ID QYK9M0 PRELIMINARY; PRT; 1150 AA.
AC QYK9M0;
DT 01-OCT-2000 (Tremblrel. 15, Created)

```

```
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
GN PYCA OR BH2625.
OS Bacillus halodurans.
OC Bacteria: Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirata C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL; AP001516; BAB06344.1; -.
DR HSSP; P24182; 1BNC.
DR InterPro; IPR000089; Biotin_lipoyl.
DR InterPro; IPR000901; CPSase.
DR InterPro; IPR000891; HMGL-like.
DR InterPro; IPR003379; PYC_OADA.
DR Pfam; PF02785; Biotin_carb_C; 1.
DR Pfam; PF00364; biotin_lipoyl; 1.
DR Pfam; PF00289; CPSase_L_chain; 1.
DR Pfam; PF02786; CPSase_L_D2; 1.
DR Pfam; PF00682; HMGL-like; 1.
DR Pfam; PF02436; PYC_OADA; 1.
DR TIGRFAMS; TIGR01235; pyruv_carbox; 1.
DR PROSITE; PS00867; CPSASE_2; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 1150 AA; 128858 MW; ED6788BB8A9F3BD4 CRC64;

Query Match 43.2%; Score 2535.5; DB 16; Length 1150;
Best Local Similarity 46.9%; Pred. No. 2.4e-127;
Matches 540; Conservative 185; Mismatches 399; Indels 27; Gaps 13;

QY 26 LPAFKILVANRGEIAVRAFALEATGATVAIYPREDGSRFSRASEAVRIGTEGSPV 85
DB 4 LKNIKVLVANRGEIAIRIFRACTELHIRTVAIYSDGTGAYHRYKADAEAYLVGEGKKPI 63

QY 86 KAYLDEIETGAIAKKYKADAIYPGYFLSENAQLARECAENGITFGPTPEVLDTGDKS 145
DB 64 EAYLDIEGIIETIAKRGVDAIHGPGYFLSENFIEFAKRCHEEGIIIFGPELHLVMEGDKV 123

QY 146 RAVTAAKKAGLPVLAEST-PSKNIDDIVKSAEGQTPFIFVKAVAGGGGGRMRVSSPEL 204
DB 124 QAREQAIKANLPVIGSDGVSLEDDYKAFADKHGYPFIITKAALGGGGGRMIVRSENDV 183

QY 205 RKLATEASREAAAGDGVSVYVRAVINPOHIEVQILGDRTEGVVHLRYDRDCSLORRHOK 264
DB 184 QESYERAKSEAKAFAFNDEYVVEKFTENPKHIEVQILADKHGNTLHLRYDRDCSVQRHOK 243

QY 265 VVEIAPAHLDELDRICADAVKFCRSIGYOGAGTVEFLVDEKGNHVFTEMNPRIQVEH 324
DB 244 VVEVAPSVLSDEVRERICAAVQLAENVNVVAGTVFELVDREGNFYFIEVNPRIQVEH 303

QY 325 TVTEEVTEVDLVKAQMRLAAGATL--KEGL--TQDKIKTHGAALQCRITTEDPNNGFRPD 381
DB 304 TITETWTGIDIVGSQLFDGEHLHGRIPKOEIVCHGYAIQSRVTTEDPSNGFLPD 363

QY 382 TGTITAYRSPGGAGVRLD--CAAQLGGEITAHFDSMLVKMTCRSGDPETAVARAQALAEF 440
DB 364 TGRINAYRSGGGFVRLDAGNGFQGAIVTIPYDLSLVKYSTWALTTEGAAKMLRNREF 423

QY 441 TVSGVATNTGFRALLREEDFTSKRTATGFIGDHPHLLQAPADDEQGRILDYLDVTVN 500
DB 424 RIRGIKNTIAFLENNVQHROFLSGEYNTSFIDQTPELFVFPKRKDRGCKMLSFIGETIVN 483

QY 501 KPHGVPRKQVAAPIDK--LPNIK-DLPLPRGSRDLKQLGPAFAFARDLREQDALAVTDTT 557
DB 501 KPHGVPRKQVAAPIDK--LPNIK-DLPLPRGSRDLKQLGPAFAFARDLREQDALAVTDTT 557
```

```
DB 484 GYPGLE-KTKKPVFDKPPVKLKLSEIPDGTQKILDOHGLEGLAKWVKQKVLVLLTDTT 542
QY 558 FRDAHOSLLATRVRSFALKPAAEAVALKLTPELLSVEAWGATYDVANRFLFEDPWRDLDE 617
DB 543 FRDAHOSLLATRVTRHDLKQIAEPTARLLPNLFSMEMMGATFDVAMRFLHEDPWERLLI 602
QY 618 LREAMPNVNIQMLLRGRNTVGYTPYDSCVCRAFVKEAASGVDFIRFDALNDVDSMRPA 677
DB 603 LRKKAPNVLFQMLLRASNAVGYKNYPDNLIRFVYDKSANAGIDVFRIFDSNLNVEGKLA 662
QY 678 IDAVLETNTAVAEVAMAYSGDSDPNKLYTLDYLLKMAEBEIVKSGAHILAIKDMAGLLR 737
DB 663 IEAVGEAN-KIAEATICYTGDIILSSRPKYDLAYYKKLAKELAAGAHILGIDMAGLLK 721
QY 738 PAATVTLVTALRRFEDLPVHVHTHDAGGOLATYFAAAQAGADAVDCASAPLSGTTTQPS 797
DB 722 PEAYQLVLAELKDTVTIPVHLHDTHTSGNGIFTYARAIEAGVDIVDVAVSSMAGLTQPS 781
QY 798 LSAIVAAFAHTRRDGTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGQL 857
DB 782 ANSLYVALADSERQPNVNITALEQLAEFWEETRKFYAGFESGMNAPHTVEYHEMPGQY 841
QY 858 SNLRQAATAGLADRFELIEDNYAANVEMLGRPTKVTTPSSKVVGDALHLVAGVDPADF 917
DB 842 SNLQQAQAKAVGLGRWNEVKMYRTVNDMFGDVYKVTTPSSKVVGDALYVQNDLTBEV 901
QY 918 AADPKYDIPDSVIAFLRGLGNPPGCGPEPLRTRALEGRSEKAPLTVPEEEQAHLD 977
DB 902 YENGHKLDFFDPSVVEFEGQLGQYQGFKKLOEIIILKGRK---PITNRPGENMEPIQF 957
QY 978 DSKER-RNSLNR-----LLFPKPTTEEFLEHRRFRFGNTSALDRDFEYFGLVGRE 1026
DB 958 EAIKEELYNKLDQVTSHDILSYALPKVFMFEFEPQTGFDVSLDTPTFYFGLRPGEE 1017
QY 1027 TLRLPDVRLPLVRLDAISEPDDKGMNVNVANVNGQIRPMRVDRDSVESVTAEKADS 1086
DB 1018 IEVEIEQGKT-LIVKFISLSKPDQDGNRIYVYFELNGPREVLINKDQSVKTSIISRPKADK 1076
QY 1087 SNKGHVAAPFAG-VVTVTVAGDEVKAGDAVAIIIEAMKMEATITASVDGKIERYVVVPAAT 1145
DB 1077 SNPNHIGASPGVTVVVALVEKDKVKQGDHLMTEAMKMETTVQAPDFGEVVALHVKGDG 1136
QY 1146 KVEGGDLIVVV 1156
DB 1137 AIQTGDLLEIV 1147

RESULT 8
Q9KW04 PRELIMINARY; PRT; 1148 AA.
AC Q9KW04;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Pyruvate carboxylase (EC 6.4.1.1).
GN PYCA.
OS Bacillus subtilis.
OC Bacteria: Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entlian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Chim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
```



```

RX MEDLINE=21100417; PubMed=11167010;
RA Yoder J.A., Litman G.W.;
RT "The zebrafish ftlh, scl3a2, men1, pc, fgf3 and cyd1 genes define two
RT regions of conserved syntenic linkage group 7 and human
RT chromosome 11q13.1";
RL Gene 261:235-242(2000).
CC -!- COFACTOR: BIOTIN (BY SIMILARITY).
DR EMBL; AF295372; AAG37836.1; -.
DR HSSP; P24182; 1BNC.
DR InterPro; IPR001882; Biotin_attach.
DR InterPro; IPR000089; Biotin_lipoyl.
DR InterPro; IPR000901; CPSase.
DR InterPro; IPR000891; HMG-Like.
DR InterPro; IPR003379; PYC_OAD.
DR Pfam; PF02785; Biotin_carb.C; 1.
DR Pfam; PF00364; Biotin_lipoyl; 1.
DR Pfam; PF00289; CPSase_L_chain; 1.
DR Pfam; PF02786; CPSase_L_D2; 1.
DR Pfam; PF00682; HMG-Like; 1.
DR Pfam; PF02436; PYC_OAD; 1.
DR TIGRfams; TIGR01235; pyruv_carbox; 1.
DR PROSITE; PS00188; BIOTIN; 1.
KW Biotin.
SQ SEQUENCE 1180 AA; 129884 MW; 26118F2691400E24 CRC64;

Query Match 43.0%; Score 2523.5; DB 13; Length 1180;
Best Local Similarity 46.6%; Pred. No. 1.le-126;
Matches 550; Conservative 179; Mismatches 419; Indels 31; Gaps 16;

QY 6 LGLLLLGIITLV-----SPTSS-TLP--AFKKILVANGEIAVAFRAALETGAA 54
DB 5 LGSRAKTGILLIRRAAGLFCRSASLTSQTLEYPHPKIKVMVANRGEIAIRVFACTELGIR 64

QY 55 TVAIYPREDGSGPHRSFASFAVRIGTEGSPVKAYLDIDEIIGAARKVKADAIYPGGFLS 114
DB 65 TVAVISEQDTGQMHKQADEAYLIGRLSPVAAYLHPDIIKVKAKENNVDAIHPGYGFLS 124

QY 115 ENAQLARECAENGITIGTPEVLDITGDKSRVAVTAKKAGLPVL-AESTPPSKNIDIVK 173
DB 125 ERADFAQAACAEAGVRIGFSPSEVVRKMGDKVEARALAKAGVPVPGTADPISCLQEAQE 184

QY 174 SAGQYPIPIKAVAGGGGMRFRVSDDELKRLATEAREAAAFGDSGVYVERAVINP 233
DB 185 FAKTYDPIIFKAAAGGGGMRVRYNEELENYORAYSEALAAAFNGALFVEKFTIEP 244

QY 234 QHTEVOILGDRTEGVHLVERDCSLORRHOKVVEIAPAOHLDPDLRDRICADAVKFCRSI 293
DB 245 RHIEVOILGDKYGNVJHLYERDCSIQRRHOKVVEIAPAOHLDSHLDRDLTLDVSNLAKQV 304

QY 294 GYGAGTVFELVDEKGNHVFIEKNPRIQVHTVEVTEVDLVKAQMRLAAGATLRELGL 353
DB 305 GYENACTVEFLVDKHKHYFIEVNSRLQVEHTVEITDVLVHAQLRVCEGRSLPELGL 364

QY 354 TDKIKTHGAALQCRITTPDNNPFRPDGTITAYRSPGAGVRLDGAAL-QGEITAHF 412
DB 365 EODKIQINGCAIQCRVITDPDPSRGFQDGTGRIEVRSGEGMIRLSDSAFAFGAIIISPHY 424

QY 413 DMLVMTCTGSGDFETAVARAQALAEFFVSGVATNIGFLRALLREEDFTSKRIATGFIG 472
DB 425 DSVLVKVIASGKDLPTAATKMRHALTEFRVGVKTNIPFLQNVLSQSLYATFVDTFQFD 484

QY 473 DIPHLQAAPADDEQGRILDYLAADVTVNKPVGVRP-KDVAAPIDK-LPNIKDLPPLRGRSR 530
DB 485 ENQNLENKPTQNAOKLHLYGHVMVNGPMTPIPVAKPSPVDPVIPSVSLSGEPPLGR 544

QY 531 DRKLQGLFAAFARDLEQDALAVDTTFRDAHQSLLATRVRSFALKPAAEAAVAKLTPPELL 590
DB 545 EVLLREGPEGFARAIRHQHGLLLMDTFRDAHQSLLATRVTHDLAKKIAPYVSHNSNLF 604

QY 591 SVEAMGGATYDVAMRFLFDPDWRDLDELREAMPNVNIQMLLGRNTVGYTPYVSDVCRFA 650
DB 605 SLENMGATFADAMRFLSECPNKRQLQELRALIPNVFQMLLRGANAVGVTNYPDNVAFK 664

```

```

QY 651 VKEAASSGVDFRIFDALNDVSOVRPAIDAVLETTAVAEVAMAYSGDLSDPNEKLYTLD 710
DB 665 CEVAKENGMDIFRVFSLNYPNMLGMEAGAAG-GVVEAAISYTGDSVDPWRQKYSLD 723
QY 711 YLLKMAEPIVKSGAHILAIKDMAGLLRPAAVTKLVLTAREF-DLPVHVHTHTAGGOLA 769
DB 724 YLLKLADELKAGTHILSKIDMAGLLKPOASRLLEALRDRFPDIPIHVHTHTAGAGVA 783
QY 770 TYFAAAGAGADAVDASAPLSGTSOPSLSAIVAAFAHTRDTCGLSLEANSYLEPYWEAV 829
DB 784 AMLACAQAGADIIVDAVDSMAGMTSQPSMGAIVACTKGTLDITGSLDKVDFDYSEYWEVA 843
QY 830 RGLYLPFESGTPGTGR--VYRHEIPGQSLNRAQATALGCLADRFELIEDNYAAVNEML 887
DB 844 RGLYAPFDCATMKSGNADVYENIEPGQVYTNLHFQAHSMGLGNKFVEVKKATKANLL 903
QY 888 GRPTKVTPTSSKVVGDIALHLVAGVDPADFAADPOKYDIPDSVIAFLRGEIKNPPGWPE 947
DB 904 GDLIKVTPSSKIVGDLAQFMV-QNLSRAEVEKRADELSFPLSVVEFLQGHIGHPGGFPE 962
QY 948 PLRTRALEG--RSEGRAPLTEVPPEEQAHLDA-----DSKERNSLNRLLEPKPTEEF 999
DB 963 PFRSKVLKSLPRIEGR-PGASLPMPDFEALSGLRANGHDEITPEDVMSAAMPKVFQEF 1021
QY 1000 LEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRPLLVRLDAISEPDDKGMNRYVAN 1059
DB 1022 KEFTSTFGPVDCLNTRFLDGPKEAEFFQVELEGRKI-LHIKALAGDLNKSQREYFFE 1080
QY 1060 VNGQIRPMRVDRSVSVTATAEKADSSNKGHVAAAPAG-VVTVTVAEGDEVKAGDAVAI 1118
DB 1081 LNQGLRSVLVKDTAAAMKEMHFHPKALKDVRGQVGPAPGPKVVEVKVAKGQKVEKQPLCV 1140
QY 1119 IEAMKWEATITASVDGKIERVVVPAATKVEGGDLIVVVS 1157
DB 1141 LSAKMETVTVNSPIGSIISKVHVNADSLSEGEDLILEIT 1179

RESULT 10
P94448 PRELIMINARY; PRT; 1147 AA.
AC P94448;
DT 01-MAY-1997 (TEMBLrel. 03, Created)
DT 01-MAY-1997 (TEMBLrel. 03, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE Pyruvate carboxylase (EC 6.4.1.1).(pyruvic carboxylase) (PCB).
OS Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Geobacillus.
OX NCBI_TaxID=1422;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1041;
RX MEDLINE=97354293; PubMed=9210587;
RA Kondo H., Kazuta Y., Saito A., Fujii K.-I.;
RT "Cloning and nucleotide sequence of Bacillus stearothermophilus
RT pyruvate carboxylase.";
RL Gene 191:47-50(1997).
CC -!- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
CC INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
CC ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE CARBOXYL
CC GROUP TO PYRUVATE IN THE SECOND. CATALYZES THE INITIAL REACTIONS
CC OF GLUCOSE SYNTHESIS FROM PYRUVATE (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: ATP + PYRUVATE + HCO(3) (-) = ADP +
CC ORTHOPHOSPHATE + OXALOACETATE.
CC -!- COFACTOR: BIOTIN (BY SIMILARITY).
CC -!- PATHWAY: GLUCONEOGENESIS.
CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -!- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, PARTICULARLY FROM
CC EUKARYOTES, LIPOAMIDE TRANSFERASES AND CARBAMYL PHOSPHATE
CC SYNTHETASES.
DR EMBL; D83706; BAA12072.1; -.
DR HSSP; P24182; 1DVL.

```

[illegible]

Qy	959	E-GKAP-----LTEVEPEEQAHLDADDSKKERNLSNRLLPKPKTPEEFLEHRR	1005
Db	985	KLDKRLGLYLEPLDLAKIKSQIRENYGAATYDVA-----SYAWPKVPEDYKKFVAK	1037
Qy	1006	FGNTSALDDREFFYGLVEGRETLIRLPDVRTPLLVRLDAISE-PDDKGRNRVWVANVGOI	1064
Db	1038	FGDLSVLPRFYFLAKPEIGEETHVELEKGV-LIUKLAIAGLSEQTQGREVFYEVNGEV	1096
Qy	1065	RPNRVRDRSVESVATAEAKDSNKGHVAAPFAGVVT-VTVAEGDGVKAGDAVAIEAMK	1123
Db	1097	ROYSVDDKKASVENTARPAEGLGDSQVGAPMSGVVVEIRVHDLGVKKGDPIAVLSAMK	1156
Qy	1124	MEATITASVDGKIERYVVVPAATKVEGGDLI	1153
Db	1157	MEMVISAPHSGKVSSLLVLEKGDSDVDQDLV	1186
RESULT	12		
ID	Q8Y846	PRELIMINARY;	PRT: 1146 AA.
AC	Q8Y846;		
DT	01-MAR-2002 (TrEMBLrel. 20, Created)		
DT	01-MAR-2002 (TrEMBLrel. 20, Last sequence update)		
DE	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)		
DE	pyCA protein.		
GN	pyCA OR LM01072.		
OS	Listeria monocytogenes.		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;		
OC	Listeriaceae; Listeria.		
OX	NCBI_TaxID=1639;		
RP	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=EGD-E / SEROVAR 1/2a;		
RX	MEDLINE=21537279; PubMed=11679669;		
RA	Glaser P., Frangeul P., Buchreiter C., Rusnlok C., Anand A.,		
RA	Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,		
RA	Charbit A., Chetoui F., Couve E., de Daruvar A., Dehoux P.,		
RA	Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dusauguet O.,		
RA	Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,		
RA	Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,		
RA	Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapat G.,		
RA	Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,		
RA	Nordisk G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,		
RA	Rommel B., Rose M., Schlueter T., Simoes N., Tierrez A.,		
RA	Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;		
RT	"Comparative genomics of Listeria species.";		
RL	Science 294:849-852(2001).		
DR	EMBL; AL591977; CAC99150.1; -.		
DR	ListiList; LM001072; -.		
DR	InterPro; IPR000089; Biotin_lipoYL.		
DR	InterPro; IPR000901; CPSase.		
DR	InterPro; IPR000891; HMGL-like.		
DR	InterPro; IPR003379; PYC.OADA.		
DR	Pfam; PF02785; Biotin_carb.C; 1.		
DR	Pfam; PF00364; biotin_lipoYL; 1.		
DR	Pfam; PF00289; CPSase_L_chain; 1.		
DR	Pfam; PF02786; CPSase_L_D2; 1.		
DR	Pfam; PF00682; HMGL-like; 1.		
DR	Pfam; PF02436; PYC.OADA; 1.		
DR	TIGRFAMS; TIGR01235; pyrUV_carbox; 1.		
DR	PROSITE; PS00866; CPSASE_1; 1.		
DR	PROSITE; PS00867; CPSASE_2; UNKNOWN_1.		
QW	complete proteome.		
QW	SEQUENCE 1146 AA; 128049 MW; A062F88CA092B6F CRC64; -		

Query Match 42.5%; Score 2494; DB 16; Length 1146;
Best Local Similarity 45.7%; Pred. No. 4e-123;
Matches 525; Conservative 186; Mismatches 408; Indels 30; Gaps 11;

Qy 30 KKIIVANRGSIATVAFRAALLETGAATVAIYIPREDRSGSFHRSEAFSEAVRICETGSPVKAYL 89
|||||
Db 5 KKIIVANRGSIATVAFRAALLETGAATVAIYIPREDRSGSFHRSEAFSEAVRICETGSPVKAYL 64
|||||

```
QY 90 DIDEIIGAAGKADAIYPGYGFLSENAQLARECAENGITFIGTPBEVLDTGDKSRVAT 149
DB 65 DIENIEIAKESGADAIHPGYGFLSENAQLARECAENGITFIGTPBEVLDTGDKSRVAT 124
QY 150 AAKKAGLPVLAEST-PSKNIDDIIVKSAEGQYPIFVAVAGGGGRGMRFVSSPDELKLA 208
DB 125 QALLADIPVPGSNGPVAGIKEVEFEKNGKGYPLMIKASLGSGGGRGMRFVSSPDELKLA 184
QY 209 TEASREAAAGDGSVYVERAVINPOHIEVOILGDRTEGVVHLYERDCSLORRHKVVEI 268
DB 185 ERASSEAAAGDGSVYVERAVINPOHIEVOILGDRTEGVVHLYERDCSLORRHKVVEI 244
QY 269 APAQHLDELDRICADAVPCRSIGYGAGTVEFLVDEKGNHVFIEPNRIQVHTVTE 328
DB 245 APCNAITSELNRICDAVKLMKNDVINAGTVEFLV-EGDDFFYFIEVNPVQVHTITE 303
QY 329 EVTEVDLVKAOMLAAGATKELGLT---QDKIKTHGAALQCRITTEDPNNGFRPDTGTI 385
DB 304 MITGIDIVQSOLFADGYALHQDLVAIPKQEDIHGSAIQSRITTEDPNLNNFMDPTGRV 363
QY 386 TAYRSPGAGVRLD-GAALQGEITTAHFDMSLVKMTCRGSDFFETAVARAQALAEFTVSG 444
DB 364 DTYRSTGFGVRLDAGNGFQGVVTPFYDSLLVKLTGWTGTFEQATKMRNLIEFRIG 423
QY 445 VATNIGELRALLREDEFTSKRIATGFTGDHPHLLQAPADDEQGRILDYLDVTVNPKHG 504
DB 424 VKNIPENLVNVRHDPDFASNGYNTSFIDTTPELFKPHIRDGKTLYIGNVTVNGFGP 483
QY 505 VRPKDVAAPIDKLPNIKDLNLP----LPRGSDRLKQLGPAFAFARDLREODALAVTDITFD 560
DB 484 IKHRD--KPVYAEPLPKIPYGSQISGTKQLDAGKEGVGVWVKKQOEVLDTITLDR 541
QY 561 AHOSLLATVRSPFALKPAEAAVAKLTPELLSVEANGATYDVAMRELFDPPWRDLRE 620
DB 542 AHOSLLATVRSKDIFOVADAMAHLLPNMFSFEMWGGATFDVAIRFNEPWRVLETLRK 601
QY 621 AMPVNTOMLLRGNTVGYTPDSCVRAVKEAASGVDFIRIFDALNDVSMRPAIDA 680
DB 602 QIPNVMQMLLRGANAVKYNYPDNVIREFVKOSAGSDVDFRVFDSLNKIKOMEVSIDA 661
QY 681 VLENTAVAEVAMAYSGLSDPNEKLTLDYLYLKMAEIEVKSGAHLAKDMAGLLRPAA 740
DB 662 VREAG-KVSEAAICYTGIDIDDDTTRKYTIDYKDMAKELVAQGTGHLGKDMAGLLKPOA 720
QY 741 VTKLVTLRREFDLPVHVHTHTAGGOLATYFAAAQAGADAVDCAAPLSGTTTSPQSLSA 800
DB 721 AYRLIGELKDTVDVPIHLHTHTDSGNGIYTYAAAVSAGVDIVDVASSAMSGATSPQSMTG 780
QY 801 IVAFAHTRRDTGLSLSEAVSDLEPYWEAVRGLYLPPESGTPGPTGRVYRHEIPGGOLSNL 860
DB 781 LYGLVNGNRQTNLDAQNSQIINHVEDVRHYKDFDNALNSPQTEVYIHEMPGGQYTNL 840
QY 861 RAQATLGLADREFELIEDNTAAVNEMLGRPTKVTTPSSKVVGDLLALHLVAGVDPADFAAD 920
DB 841 QQAIAVGLGDRWDEYKEMTYVNVQMGDIVKVTTPSSKVVGDLLALPMVQNELSEEDVYEK 900
QY 921 POKYDIPDSVIAFLRGLGNPPGMPPELTRALEGRSEKAPLITEVPPEEQALHDADS 980
DB 901 GDTIDPDSVIEEFMGIEGQYGGFPEKLOKLVKGT-----PLTDPCALMEPVNFVDV 956
QY 981 K-----ERRNSNRLPLPTEPELEHRRRFGNTSALDREFFYGLVEGRETLI 1029
DB 957 KAELEKMGVEYTEKDVISILYKPVFLDYQDMINKYGDVTLDTFTFYKGRMLGETIEV 1016
QY 1030 RLPDVRTPLVLRLDAISEPDDKGMNVNANVNGQIRPMVRVSRESVATAEKADSSNK 1089
DB 1017 ELEKGGK-LLIKLSNGEPTADTRIVYFELNGOPREINIQDMNVQSTVIARRKIDTTNP 1075
QY 1090 GHVAAPFAG-VVTVTVAGDEVKAGDAVATIEAMKMEATITASVDGKIERVVVPAATKVE 1148
DB 1076 ERVGATMTGSVIOVVVVKSGDKVKKGDPPLITIAMKMETTIQAPFDEGVSSIVVSGDGTIE 1135
QY 1149 GGLIVVVS 1157
```

```
DB 1136 SGDILLIEVN 1144
RESULT 13
Q92CW1 PRELIMINARY; PRT: 1146 AA.
AC Q92CW1;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE PYCA protein.
GN PYCA OR LIN1060.
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Listeriaceae; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CLIP 11262 / SEROVAR 6A;
RX PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
BAQUERO F., Berche P., Blocker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetoui F., Couve E., de Daruvar A., Denoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fshih H., Garcia-del Portillo F., Garrido F.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain F., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Klett J., Kuhn M., Kunst F., Nedjari H.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsted B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL; AL596167; CAC96291.1; -
DR ListList; LIN01060; -
DR InterPro; IPR000089; Biotin_lipoyl.
DR InterPro; IPR000901; CPSase.
DR InterPro; IPR000891; HMGL-like.
DR InterPro; IPR003379; PYC_OADA.
DR Pfam; PF02785; Biotin-carb_C; 1.
DR Pfam; PF00364; Biotin_lipoyl; 1.
DR Pfam; PF00289; CPSase_L_chain; 1.
DR Pfam; PF02786; CPSase_L_D2; 1.
DR Pfam; PF00682; HMGL-like; 1.
DR Pfam; PF02436; PYC_OADA; 1.
DR TIGRfams; TIGR01235; pyruv_carbox; 1.
DR PROSITE; PS00866; CPSASE_1; UNKNOWN_1.
DR PROSITE; PS00867; CPSASE_2; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 1146 AA; 128035 MW; 3E23FFB4A289C60F CRC64;

Query Match 42.5%; Score 2491; DB 16; Length 1146;
Best Local Similarity 45.9%; Pred. No. 5.9e-125;
Matches 527; Conservative 183; Mismatches 409; Indels 30; Gaps 12;

QY 30 KILVANGETIAVRAALFETGAATVAIYPREDGRSFHRSFASPAVRIGTEGSPVKAYL 89
DB 5 KKLVANGETIAVRAALFETGAATVAIYQEDTGSFHYKSDIAYLVGAGKPIDAYL 64
QY 90 DIDEIIGAAGKADAIYPGYGFLSENAQLARECAENGITFIGTPBEVLDTGDKSRVAT 149
DB 65 DIENIEIAKESGADAIHPGYGFLSENAQLARECAENGITFIGTPBEVLDTGDKSRVAT 124
QY 150 AAKKAGLPVLAEST-PSKNIDDIIVKSAEGQYPIFVAVAGGGGRGMRFVSSPDELKLA 208
DB 125 QALLADIPVPGSNGPVAGIKEVEFEKNGKGYPLMIKASLGSGGGRGMRFVSSPDELKLA 184
QY 209 TEASREAAAGDGSVYVERAVINPOHIEVOILGDRTEGVVHLYERDCSLORRHKVVEI 268
DB 185 ERASSEAAAGDGSVYVERAVINPOHIEVOILGDRTEGVVHLYERDCSLORRHKVVEI 244
QY 269 APAQHLDELDRICADAVPCRSIGYGAGTVEFLVDEKGNHVFIEPNRIQVHTVTE 328
```



```

Db 245 APCNAITSLNRICDAAVKLMKNVDY INAGTVEFLV-EGDNFYFEVNPVQVEHTITE 303
QY 329 EYTEVDLVKAQRLAAGATLKGELGT--ODKIKTHGAALQCRITTEDPNNFRPDGTGI 385
Db 304 MITGDIVOSQFIADGAYALHOLVAIPKQEDIIHGSAIQSRITTEDPLNNFMPDGTGRV 363
QY 386 TAYRPGGAGVRLD-GAALGGEITAHFDSMLVKMTCRGSDSETAVARAQALAEFTVSG 444
Db 364 DTYRSTGGFVRLDAGNGFGVIVPFYDLSLVKLTGWTGTEQATKRRNRLIERIRG 423
QY 445 VATNTGFLRALLREEDFTSKRIATGIGDHPHLLQAPPDQSGRILDYADVYVKNKPH 504
Db 424 VKTNPFLNVRHPDFAFSGNNTSPIDTTPLEFKPHIRDRGKTLRIGNVTVNGFGG 483
QY 505 VPKQVAAPIDKLPNIKDL-P---LPRGSRDLKQLGPAAFARDLEQDALAVTDTTFRD 560
Db 484 IKHRD--KPYAEPRLPKIPYGSQIAPGTQKILDAKPGGVVDWVKQBEVLLTDTFLKD 541
QY 561 AHQSLLATRVRSFALKPAAEVAKLTPELLSVKATYDVAMRFLFEDPMDRLDELRE 620
Db 542 AHQSLLATRVRSKDIQIADAMAHLLPNMFSEFMWGATFDVAYRFLNEDPWVRLTLRK 601
QY 621 AMPNVNIQMLLRNTVGTPTPDSVCRAPFVKEAASSGVVDIERIFDALNDVSOMRPAIDA 680
Db 602 QIPNVMFQMLLRGANAAGYKNTPDNVIREFVKQAQSGVDVFRVDSLNWIKGMEVSIDA 661
QY 681 VLENTTAVAEVAMVSGDLSDNKELYTLDYLYKMAEEIVKSCAHLTAIKDMAGLLRPAA 740
Db 662 VREAG-KVYEATICYGDDDDTRKTYTIDYKDMAKELVAOGTHILGDKMAGLLKPOA 720
QY 741 VTKLVTALRRFDPVHVHTHTAGQLATYFAAAQAGADAVDAGSAPLSGTTSQPSLSA 800
Db 721 AYRLIGELKDTVDVPIHLHTHTDSGNGIVTYAAAVSAGVDIVDVASSAMGATSQPSMTG 780
QY 801 IVAFAHTTRDTGLSLEAVSDLEPTWEAVRGLYLPFESGTGPTGRVYRHEITPGGOLSNL 860
Db 781 LYYGLVNGNRQTNLDQNSQIINHMYEDVRHYKQDNALNSPQTEVYIHEMPGGQYTNL 840
QY 861 RAQATGALGLADRFELTDNYAAVNMELGRPTKVTTPSKVVGDLALHLVAGVDPADFAAD 920
Db 841 QQAATAVGLGDRWDEKEMVTYVQNFQDIVKVTTPSKVVGDLALFWQNELTEEDVYEK 900
QY 921 POKYDIPDSVIAFLRGELOPNPGWPEPLTRALEGRSEKAPLTEVP-----EEE 971
Db 901 GDTIDPDSVIEFFMGEIGOPYGGPEKQLKLVKGR-----PLADRPCALMEPVNAEV 956
QY 972 QAHLDADDSKE--RRNSLNRLFPKPTFEFLHRRRFGNTSALDDREFPYGLVEGRETLI 1029
Db 957 KAELEKMGYEPSEKDVISYILYKPFVDYQEMISKYGDVTVDLTPFYKMRGLGETIEV 1016
QY 1030 RLDPVTRTLVRLDAISEPDDKGRNVVANVNGOIRPMVRDRSVESVTATAEKADSSNK 1089
Db 1017 ELEKGI-LLIKLSNGEPIADGTRVYIFELNGQPREINIQDMNVGUSTVIARRKIDTTNP 1075
QY 1090 GHVAAPFAG-VYTVTVAGDEVKAGDAVIAIEAMKMEATITASVDGKIERVVVPAATKVE 1148
Db 1076 EHVGTMTGTSVIVVVKKGSVKKGPDLITTEAMKMETTIQAPFDGEVSSIVVSDGDTIE 1135
QY 1149 GGDLIWVS 1157
Db 1136 SGDLLIEVN 1144

```

RESULT 14

```

Q97FR7 ID Q97FR7 PRELIMINARY; PRT: 1144 AA.
AC Q97FR7;
DT 01-OCT-2001 (TReMBLrel. 18, Created)
DT 01-OCT-2001 (TReMBLrel. 18, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Pyruvate carboxylase, PYKA.
GN CAC3660.

```

```

OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I., Daly M.J.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL: AE007763; RAK80607.1; -.
DR InterPro: IPR001882; Biotin_attach.
DR InterPro: IPR000089; Biotin_lipoyl.
DR InterPro: IPR000901; CPSase.
DR InterPro: IPR000891; HMGL-like.
DR InterPro: IPR003141; PHP_N.
DR InterPro: IPR003379; PYC_OADA.
DR Pfam: PF02785; Biotin_carb_C; 1.
DR Pfam: PF00364; biotin_lipoyl; 1.
DR Pfam: PF00289; CPSase_L_chain; 1.
DR Pfam: PF02786; CPSase_L_D2; 1.
DR Pfam: PF00682; HMGL-like; 1.
DR Pfam: PF02436; PYC_OADA; 1.
DR PRINTS: PR00098; CPSASE.
DR SMART: SM00481; POLIITAC; 1.
DR TIGRFAMs: TIGR01235; pyruv_carbox; 1.
DR PROSITE: PS00188; BIOTIN; UNKNOWN_1.
DR PROSITE: PS00866; CPSASE_1; 1.
DR PROSITE: PS00867; CPSASE_2; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 1144 AA; 127709 MW; 519FA29A8008F326 CRC64;

```

```

Query Match 42.5%; Score 2490.5; DB 16; Length 1144;
Best Local Similarity 45.2%; Pred. No. 6.2e-125;
Matches 519; Conservative 204; Mismatches 392; Indels 33; Gaps 14;

QY 29 FKKILVANGETAVRAFALETCATVAIYPRDGRSGHRSFASAVRIGTSGSPVKAY 88
Db 5 FKKVLVANGETAIRIFRACHELGITVAIYSEEDLALFRTRADESYLIGQNRKGPVDA 64
QY 89 LDTEIIGAACKVKADAIYPGYGFLSENAOLARECEANGITFGPTPEVLDTGDKSRV 148
Db 65 LNIDEIINALKGVDAIHPGYGFLSENSEFSRRCTEAGIEFGPTGDMMDKLGDKINSK 124
QY 149 TAAKAGLPVL-AESTPSKNIDDIVKSAEQTYPIFVAVAGGGGGRMFVSSPDELRLK 207
Db 125 LAAKAAGVKTIPGVKPIETEQAIEFARTCGYPVMVKAAGGGGRMIRIVEEDLIAA 184
QY 208 ATEASREAAAFGGDGVYVERAVINPOHTEVQILGORTGEVHLYERDCSLQRHOKVE 267
Db 185 CRSASEAKKAFGIEDIFIEKYLEGPKHIEVQVLGKYGNIVHLYERDCSVQRHOKVIE 244
QY 268 IAPQAHLDPDLRICADAVKFCRSIGYOGAGTVEFLVDEKGNHVFTEMPPRIQVEHTVT 327
Db 245 LTPAVSMSEKRLCEICEDALKIARSTGYRSAGTLEFLDKHGNHYFTEMPPRQVEHTIT 304
QY 328 EEVTEVDLVKAQRLAAGATLK--ELGL-TQDKIKTHGAALQCRITTEDPNNFRPDGTG 384
Db 305 EMVTGIDIVSQILIAEGYKLSNPSVQINSQEDIHVNGYAIQCRITTEDPSNSFAPDTGK 364
QY 385 ITAYRSPGGAGVRLDGA-AQLGGEITAHFDSMLVKMTCRGSDSETAVARAQALAEFTVS 443
Db 365 IDVYRTGSGFGIRLDGGNGFTGAVISPYDLSLVKSTWSRSTFEDAIRKAIRAKETIYS 424
QY 444 GVATNIGFLRALLREEDFTSKRIATGIGDHPHLLQAPPDQSGRILDYADVYVKNKPH 503
Db 425 GVKTNIDFLNVLNHNHETFRKGLCDTNFIANNPELFEITPRIDELRVLFKIGERKVNETH 484

```


